

# Beyond the Dirac phase factor: Dynamical Quantum Phase-Nonlocalities in the Schrödinger Picture

Konstantinos Mouloupoulos\*

*University of Cyprus, Department of Physics, 1678 Nicosia, Cyprus*

(August 30, 2011)

## Abstract

Generalized solutions of the standard gauge transformation equations are presented and discussed in physical terms. They go beyond the usual Dirac phase factors and they exhibit nonlocal quantal behavior, with the well-known Relativistic Causality of classical fields affecting directly the *phases* of wavefunctions in the Schrödinger Picture. These nonlocal phase behaviors, apparently overlooked in path-integral approaches, give a natural account of the dynamical nonlocality character of the various (even static) Aharonov-Bohm phenomena, while at the same time they seem to respect Causality. For particles passing through nonvanishing magnetic or electric fields they lead to cancellations of Aharonov-Bohm phases at the observation point, generalizing earlier semiclassical experimental observations (of Werner & Brill) to delocalized (spread-out) quantum states. This leads to a correction of previously unnoticed sign-errors in the literature, and to a natural explanation of the deeper reason why certain time-dependent semiclassical arguments are consistent with static results in purely quantal Aharonov-Bohm configurations. These nonlocalities also provide a remedy for misleading results propagating in the literature (concerning an uncritical use of Dirac phase factors, that persists since the time of Feynman's work on path integrals). They are shown to conspire in such a way as to exactly cancel the instantaneous Aharonov-Bohm phase and recover Relativistic Causality in earlier “paradoxes” (such as the van Kampen thought-experiment), and to also complete Peshkin's discussion of the electric Aharonov-Bohm effect in a *causal* manner. The present formulation offers a direct way to address time-dependent single-*vs* double-slit experiments and the associated causal issues – issues that have recently attracted attention, with respect to the inability of current theories to address them.

**KEYWORDS:** Aharonov-Bohm, Gauge Transformations, Dirac Phase Factor, Quantum Phases

---

\*Electronic address: cos@ucy.ac.cy

## I. INTRODUCTION

The Dirac phase factor — with a phase containing spatial or temporal integrals of potentials (of the general form  $\int_{\mathbf{r}}^{\mathbf{r}} \mathbf{A} \cdot d\mathbf{x}' - c \int_t^t \phi dt'$ ) — is the standard and widely used solution of the gauge transformation equations of Electrodynamics (with  $\mathbf{A}$  and  $\phi$  vector and scalar potentials respectively). In a quantum mechanical context, it connects wavefunctions of two systems (with different potentials) that experience the same classical fields at the observation point  $(\mathbf{r}, t)$ , the two more frequently discussed cases being: either systems that are completely gauge-equivalent (a trivial case with no physical consequences), or systems that exhibit phenomena of the Aharonov-Bohm type (magnetic or electric)[1] — and then this Dirac phase has nontrivial observable consequences (mathematically, this being due to the fact that the corresponding “gauge function” is now multiple-valued). In the above two cases, the classical fields experienced by the two (mapped) systems are equal *at every point* of the accessible spacetime region. However, it has not been widely realized that the gauge transformation equations, viewed in a more general context, can have *more general solutions* than simple Dirac phases, and these lead to wavefunction-*phase-nonlocalities* that have been widely overlooked and that seem to have important physical consequences. These nonlocal solutions are applicable to cases where the two systems are allowed to experience *different fields* at spacetime points (or regions) that are *remote* to (and do *not* contain) the observation point  $(\mathbf{r}, t)$  (these regions being *physically accessible* to the particle, unlike genuine Aharonov-Bohm cases). In this article we rigorously show the existence of these generalized solutions, demonstrate them in simple physical examples, and fully explore their structure, presenting cases (and closed analytical results for the wavefunction-phases) that actually connect (or map) two quantal systems that are **neither physically equivalent nor of the usual Aharonov-Bohm type**. We also fully investigate the consequences of these generalized (*nonlocal*) influences (on wavefunction-phases) and find them to be numerous and important; we actually find them to be of a different type in static and in time-dependent field-configurations (and in the latter cases we show that they lead to Relativistically *causal* behaviors, that apparently resolve earlier “paradoxes” arising in the literature from the use of standard Dirac phase factors). The nonlocal phase behaviors discussed in the present work may be viewed as a justification for the (recently emphasized[2]) terminology of “dynamical

nonlocalities” associated with all Aharonov-Bohm effects (even static ones), although in our approach these nonlocalities seem to also respect Causality (without the need to independently invoke the Uncertainty Principle) – and, to the best of our knowledge, this is the first theoretical picture with such characteristics.

In order to introduce some background and further motivation for this article let us first remind the reader of a very basic property that will be central to everything that follows, which however is usually taken to be valid only in a restricted context (but is actually more general than often realized). This property is a simple (U(1)) phase-mapping between quantum systems, and is usually taken in the context of gauge transformations, ordinary or singular; here, however, it will appear in a more general framework, hence the importance of reminding of its independent, basic and more general origin. We begin by recalling that, if  $\Psi_1(\mathbf{r}, t)$  and  $\Psi_2(\mathbf{r}, t)$  are solutions of the time-dependent Schrödinger (or Dirac) equation for a quantum particle of charge  $q$  that moves (as a test particle) in two distinct sets of (predetermined and classical) vector and scalar potentials  $(\mathbf{A}_1, \phi_1)$  and  $(\mathbf{A}_2, \phi_2)$ , that are generally spatially- and temporally-dependent [and such that, at the spacetime point of observation  $(\mathbf{r}, t)$ , the magnetic and electric fields are the same in the two systems], then we have the following formal connection between the solutions (wavefunctions) of the two systems

$$\Psi_2(\mathbf{r}, t) = e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r}, t)}\Psi_1(\mathbf{r}, t), \quad (1)$$

with the function  $\Lambda(\mathbf{r}, t)$  required to satisfy

$$\nabla\Lambda(\mathbf{r}, t) = \mathbf{A}_2(\mathbf{r}, t) - \mathbf{A}_1(\mathbf{r}, t) \quad and \quad -\frac{1}{c}\frac{\partial\Lambda(\mathbf{r}, t)}{\partial t} = \phi_2(\mathbf{r}, t) - \phi_1(\mathbf{r}, t). \quad (2)$$

The above property can be immediately proven by substituting each  $\Psi_i$  into its corresponding ( $i_{th}$ ) time-dependent Schrödinger equation (namely with the set of potentials  $(\mathbf{A}_i(\mathbf{r}, t), \phi_i(\mathbf{r}, t))$ ): one can then easily see that (1) and (2) guarantee that both Schrödinger equations are indeed satisfied together (after cancellation of a few terms and then elimination of a global phase factor in system 2). [In addition, the equality of all classical fields at the observation point, namely  $\mathbf{B}_2(\mathbf{r}, t) = \nabla \times \mathbf{A}_2(\mathbf{r}, t) = \nabla \times \mathbf{A}_1(\mathbf{r}, t) = \mathbf{B}_1(\mathbf{r}, t)$  for the magnetic fields and  $\mathbf{E}_2(\mathbf{r}, t) = -\nabla\phi_2(\mathbf{r}, t) - \frac{1}{c}\frac{\partial\mathbf{A}_2(\mathbf{r}, t)}{\partial t} = -\nabla\phi_1(\mathbf{r}, t) - \frac{1}{c}\frac{\partial\mathbf{A}_1(\mathbf{r}, t)}{\partial t} = \mathbf{E}_1(\mathbf{r}, t)$  for the electric fields, is obviously consistent with all equations (2) (as is easy to see if we take

the *curl* of the 1st and the *grad* of the 2nd) – provided, at least, that  $\Lambda(\mathbf{r}, t)$  is such that interchanges of partial derivatives with respect to all spatial and temporal variables (at the point  $(\mathbf{r}, t)$ ) are allowed].

As already mentioned, the above fact is of course well-known within the framework of the theory of quantum mechanical gauge transformations (the usual case being for  $\mathbf{A}_1 = \phi_1 = 0$ , hence for a mapping from a system with no potentials); but in that framework, these transformations are supposed to connect (or map) two *physically equivalent systems* (more rigorously, this being true for ordinary gauge transformations, in which case the function  $\Lambda(\mathbf{r}, t)$ , the so-called gauge function, is unique (single-valued) in spacetime coordinates). In a formally similar manner, the above argument is also often used in the context of the so-called “singular gauge transformations”, where  $\Lambda$  is multiple-valued, but the above equality of classical fields is still imposed (at the observation point, which always lies in a physically accessible region); then the above simple phase mapping (at all points of the physically accessible spacetime region, that always and everywhere experience equal fields) leads to the standard phenomena of the Aharonov-Bohm type, where *unequal fields in physically inaccessible regions* have observable consequences. However, we should keep in mind that that above property ((1) and (2) taken together) can be *more generally valid* – and in this article we will present cases (and closed analytical results for the appropriate phase connection  $\Lambda(\mathbf{r}, t)$ ) that actually connect (or map) two systems (in the sense of (1)) that are *neither physically equivalent nor of the usual Aharonov-Bohm type*. And naturally, because of the above provision of field equalities at the observation point, it will turn out that any nonequivalence of the two systems will involve *remote* (although *physically accessible*) regions of spacetime, namely regions that do *not* contain the observation point  $(\mathbf{r}, t)$  (and in which regions, as we shall see, the classical fields experienced by the particle may be *different* in the two systems).

## II. MOTIVATION

One may wonder on the actual reasons why one should be looking for more general cases of a simple phase mapping of the type (1) between *nonequivalent* systems. To answer this, let us take a step back and first recall some simple and well-known results that originate from the above phase mapping. It is standard knowledge, for example, that, if we want

to find solutions  $\Psi(x, t)$  of the  $t$ -dependent Schrödinger (or Dirac) equation for a quantum particle (of charge  $q$ ) that moves along a (generally curved) one-dimensional (1-D) path, and in the presence (somewhere in the embedding 3-dimensional (3-D) space) of a fairly localized (and time-independent) classical magnetic flux  $\Phi$  that *does **not** pass through any point of the path*, then we formally have

$$\Psi(x, t)^{(\mathbf{A})} \sim e^{i\frac{q}{\hbar c} \int_{x_0}^x \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'} \Psi(x, t)^{(0)} \quad (3)$$

(the dummy variable  $\mathbf{r}'$  describing points along the 1-D path, and the term “formally” signifying that the above is valid *before imposition of any boundary conditions* (meaning that these are to be imposed only on  $\Psi^{(\mathbf{A})}$  and not necessarily on  $\Psi^{(0)}$ ). In (3),  $\Psi(x, t)^{(0)}$  is a formal solution of the same system in the case of absence of any potentials (hence with magnetic flux  $\Phi = 0$  everywhere in the 3-D space). The above holds because, for *all points*  $\mathbf{r}'$  of the 1-D path, the particle experiences a vector potential  $\mathbf{A}(\mathbf{r}')$  of the form  $\mathbf{A}(\mathbf{r}') = \nabla' \Lambda(\mathbf{r}')$  (since the magnetic field is  $\nabla' \times \mathbf{A}(\mathbf{r}') = 0$  for *all*  $\mathbf{r}'$ , by assumption), in combination with the above phase-mapping (with a phase  $\frac{q}{\hbar c} \Lambda(\mathbf{r})$ ) between two quantum systems, one in the presence and one in the absence of a vector potential (i.e. the potentials in (2) being  $\mathbf{A}_1 = 0$  and  $\mathbf{A}_2 = \mathbf{A}$ , together with  $\phi_2 = \phi_1 = 0$  if we decide to attribute everything to vector potentials only). In this particular system, the obvious  $\Lambda(\mathbf{r})$  that solves the above  $\mathbf{A}(\mathbf{r}) = \nabla \Lambda(\mathbf{r})$  (for **all** points of the 1-D space available to the particle) is indeed  $\Lambda(\mathbf{r}) = \Lambda(\mathbf{r}_0) + \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'$  (since  $\int_1^2 \nabla' \Lambda(\mathbf{r}') \cdot d\mathbf{r}' = \Lambda(2) - \Lambda(1)$ ), and this gives (3) (if  $\mathbf{r}$  denotes the above point  $x$  of observation and  $\mathbf{r}_0$  the arbitrary initial point  $x_0$  (both lying on the physical path), and if the constant  $\Lambda(\mathbf{r}_0)$  is taken to be zero).

What if, however, some parts of the magnetic field that comprise the magnetic flux  $\Phi$  actually *pass through* some points or a whole region (interval) of the path available to the particle? In such a case, the above standard argument is not valid (as  $\mathbf{A}$  cannot be written as a grad at any point of the interval where the magnetic field  $\nabla \times \mathbf{A} \neq 0$ ). Are there however general results that we can still write for  $\Psi(x, t)^{(\mathbf{A})}$ , if the spatial point of observation  $x$  is again outside the interval with the nonvanishing magnetic field? Or, what if in the previous problems, the magnetic flux (either remote, or partly passing through the path) is time-dependent  $\Phi(t)$ ? (In that case then, there exists in general an additional electric field  $E$  induced by Faraday’s law of Induction on points of the path, and the usual gauge transformation argument is once again not valid).

Returning to another standard (solvable) case (which is actually the “dual” or the “electric analog” of the above), if along the 1-D physical path the particle experiences only a spatially-uniform (but generally time-dependent) classical scalar potential  $\phi(t)$ , we can again formally map  $\Psi(x, t)^{(\phi)}$  to a potential-free solution  $\Psi(x, t)^{(0)}$ , through a  $\Lambda(t)$  that now solves  $-\frac{1}{c} \frac{\partial \Lambda(t)}{\partial t} = \phi(t)$ , and this gives  $\Lambda(t) = \Lambda(t_0) - c \int_{t_0}^t \phi(t') dt'$ , leading to the “electric analog” of (3), namely

$$\Psi(x, t)^{(\phi)} \sim e^{-i \frac{q}{\hbar} \int_{t_0}^t \phi(t') dt'} \Psi(x, t)^{(0)} \quad (4)$$

with obvious notation. (Notice that, for either of the two mapped systems in this problem, the electric field is zero at all points of the path). What if, however, the scalar potential has also some  $x$ -dependence along the path (that leads to an electric field (in a certain interval) that the particle passes through)? In such a case, the above standard argument is again not valid. Are there however general results that we can still write for  $\Psi(x, t)^{(\phi)}$ , if the spatial point of observation  $x$  is again outside the interval with the nonvanishing electric field?

We state here directly that this article will provide affirmative answers to questions of the type posed above, by actually giving the corresponding general results in closed analytical forms.

At this point it is also useful to briefly reconsider the earlier mentioned case, namely of a time-dependent  $\Phi(t)$  that is remote to the 1-D physical path, because in this manner we can immediately provide another motivation for the present work: this time-dependent problem is surrounded with a number of important misconceptions in the literature (the same being true about its electric analog, as we shall see): the formal solution that is usually written down for a  $\Phi(t)$  is again (3), namely the above spatial line integral of  $\mathbf{A}$ , in spite of the fact that  $\mathbf{A}$  is now  $t$ -dependent; the problem then is that, because of the first of (2),  $\Lambda$  must now have a  $t$ -dependence and, from the second of (2), there must necessarily be scalar potentials involved in the problem (which have been by force set to zero, in our pre-determined mapping between vector potentials only). Having decided to use systems that experience only vector (and not scalar) potentials, the correct solution cannot be simply a trivial  $t$ -dependent extension of (3). A corresponding error is usually made in the electric dual of the above, namely in cases that involve  $\mathbf{r}$ -dependent scalar potentials, where (4) is still erroneously used (with  $\phi(t')$  replaced by  $\phi(\mathbf{r}, t')$ ), giving an  $\mathbf{r}$ -dependent  $\Lambda$ , although this would necessarily lead to the involvement of vector potentials (through the first of (2) and the  $\mathbf{r}$ -dependence of

$\Lambda$ ) that have been neglected from the beginning – a situation (and an error) that appears, in exactly this form, in the description of the so-called electric Aharonov-Bohm effect[1, 3] as we shall see.

Speaking of errors in the literature, it might here be the perfect place to also point to the reader the most common misleading statement often made in the literature (and again, for notational simplicity, we restrict our attention to a one-dimensional system, with spatial variable  $x$ , although the statement is obviously generalizable to (and often made for systems of) higher dimensionality by properly using line integrals over arbitrary curves in space): It is usually stated [e.g. in Brown & Holland[4], see i.e. their eq. (57) applied for vanishing boost velocity  $\mathbf{v} = 0$ ] that the general gauge function that connects (through a phase factor  $e^{i\frac{q}{\hbar c}\Lambda(x,t)}$ ) the wavefunctions of a quantum system with no potentials (i.e. with a set of potentials  $(\mathbf{0}, 0)$ ) to the wavefunctions of a quantum system that moves in vector potential  $\mathbf{A}(x, t)$  and scalar potential  $\phi(x, t)$  (i.e. in a set of potentials  $(\mathbf{A}, \phi)$ ) is the obvious combination (and a natural extension) of (3) and (4), namely

$$\Lambda(x, t) = \Lambda(x_0, t_0) + \int_{x_0}^x \mathbf{A}(x', t) dx' - c \int_{t_0}^t \phi(x, t') dt', \quad (5)$$

which, however, is **incorrect** for  $x$  and  $t$  uncorrelated variables: it does **not** satisfy the standard system of gauge transformation equations

$$\nabla \Lambda(x, t) = \mathbf{A}(x, t) \quad \text{and} \quad -\frac{1}{c} \frac{\partial \Lambda(x, t)}{\partial t} = \phi(x, t). \quad (6)$$

The reader can easily see why: (i) when the  $\nabla$  operator acts on eq.(5), it gives the correct  $\mathbf{A}(x, t)$  from the 1st term, but it also gives some annoying additional nonzero quantity from the 2nd term (that survives because of the  $x$ -dependence of  $\phi$ ); hence it invalidates the first of the basic system (6). (ii) Similarly, when the  $-\frac{1}{c} \frac{\partial}{\partial t}$  operator acts on eq.(5), it gives the correct  $\phi(x, t)$  from the 2nd term, but it also gives some annoying additional nonzero quantity from the 1st term (that survives because of the  $t$ -dependence of  $\mathbf{A}$ ); hence it invalidates the second of the basic system (6). It is only when  $\mathbf{A}$  is  $t$ -independent, and  $\phi$  is spatially-independent, that eq.(5) can be correct (as the above annoying terms do not appear and the basic system is satisfied), although it is still not necessarily the most general form for  $\Lambda$ , as we shall see. [An alternative form that is also widely thought to be correct is again eq.(5), but with the variables that are not integrated over implicitly assumed to

belong to the initial point (hence a  $t_0$  replaces  $t$  in  $\mathbf{A}$ , and simultaneously an  $x_0$  replaces  $x$  in  $\phi$ ). However, one can see again that the system (6) is not satisfied (the above differential operators, when acted on  $\Lambda$ , give  $\mathbf{A}(x, t_0)$  and  $\phi(x_0, t)$ , hence not the values of the potentials at the point of observation  $(x, t)$  as they should), this not being an acceptable solution either].

What is the problem here? Or, better put, what is the deeper reason for the above inconsistencies? The short answer is the uncritical use of Dirac phase factors that come from path-integral treatments. It is indeed obvious that the form (5) that is often used in the literature (in canonical (non-path-integral) formulations where  $x$  and  $t$  are **uncorrelated** variables (and not correlated to produce a path  $x(t)$ )) *is not generally correct*, and that is one of the main points that has motivated this work. We will find *generalized results* that actually *correct* eq.(5) through extra nonlocal terms, and through the proper appearance of  $x_0$  and  $t_0$  (as in eq.(11) and eq.(12) to be found later in Section III), and these are the **exact** ones (namely the exact  $\Lambda(x, t)$ , that at the end, upon action of  $\nabla$  and  $-\frac{1}{c}\frac{\partial}{\partial t}$  satisfies exactly the basic system (6), viewed as a system of Partial Differential Equations (PDEs)). And the formulation that gives these results is generalized later in the article, for  $\Lambda(x, y)$  (in the 2-D static case) and also for  $\Lambda(x, y, t)$  (in the full dynamical 2-D case), and leads to the exact (nontrivial) forms of the phase function  $\Lambda$  that satisfy (in all cases) the system (6) – with *the direct verification (i.e. proof, by “going backwards”, that these forms are indeed the exact solutions of (6)) also being given* for the reader’s convenience. [For the “direct” and rigorous mathematical derivations see [5].]

This article gives a full exploration of issues related to the above motivating discussion, by pointing to a “practical” (and generalized) use of gauge transformation mapping techniques, that at the end lead to these generalized (and, at first sight, unexpected) solutions for the general form of  $\Lambda$ . For cases such as the ones discussed above, or even more involved ones, there still appears to exist a simple phase mapping (between two inequivalent systems), but the phase connection  $\Lambda$  seems to contain not only integrals of potentials, but also “fluxes” of the classical field-differences from *remote* spacetime regions (regions, however, that are physically accessible to the particle). The above mentioned systems are the simplest ones where these new results can be applied, but apart from this, the present investigation seems to lead to a number of nontrivial corrections of misleading (or even incorrect) reports of the above type in the literature, that are not at all marginal (and are



due to an incorrect use of a path-integral viewpoint in an otherwise canonical framework – an error that appears to have been made repeatedly since the original conception of the path-integral formulation, as we shall see). The generalized  $\Lambda$ -forms also lead to an honest resolution of earlier “paradoxes” (involving Relativistic Causality), and in some cases to a new interpretation of known semiclassical experimental observations, corrections of certain sign-errors in the literature, and nontrivial extensions of earlier semiclassical results to general (even completely delocalized) states. [As a byproduct, we will also show that – contrary to what is usually stated in earlier but also recent popular reports – the semiclassical phase picked up by classical trajectories (that are deflected by the Lorentz force) is *opposite* (and not equal) to the so-called Aharonov-Bohm phase due to the flux enclosed by the same trajectories; we will also provide 2 figures to visually assist the detailed proof of this result as well as to facilitate an elementary physical understanding of this opposite sign relation]. Most importantly, however, the new formulation seems capable of addressing causal issues in time-dependent single- *vs* double-slit experiments, an area that seems to have recently attracted attention[2, 6, 7]).

### III. 1-D DYNAMIC CASE

Let us begin with the simplest case of 1-D systems but in the most general dynamic environment, i.e. a single quantum particle of charge  $q$ , but in the presence of arbitrary (spatially nonuniform and time-dependent) vector and scalar potentials. Let us actually consider this particle moving either inside a set of potentials  $A_1(x, t)$  and  $\phi_1(x, t)$  (case 1) or inside a set of potentials  $A_2(x, t)$  and  $\phi_2(x, t)$  (case 2), and try to determine the most general gauge function  $\Lambda(x, t)$  that takes us from (maps) the wavefunctions of the particle in case 1 to those of the same particle in case 2 (meaning the usual mapping (1) between the wavefunctions of the two systems through the phase factor  $\frac{q}{\hbar c}\Lambda(x, t)$ ). [As already noted, we should keep in mind that for this mapping to be possible we *must* assume that at the point  $(x, t)$  of observation (or “measurement” of  $\Lambda$  or the wavefunction  $\Psi$ ) we have equal electric fields ( $E_i = -\nabla\phi_i - \frac{1}{c}\frac{\partial A_i}{\partial t}$ ), namely

$$-\frac{\partial\phi_2(x, t)}{\partial x} - \frac{1}{c}\frac{\partial A_2(x, t)}{\partial t} = -\frac{\partial\phi_1(x, t)}{\partial x} - \frac{1}{c}\frac{\partial A_1(x, t)}{\partial t} \quad (7)$$

(so that the  $A$ 's and  $\phi$ 's in (7) can indeed satisfy the basic system of equations (2), or equivalently, of the system of equations (10) below – as can be seen by taking the  $\frac{1}{c}\frac{\partial}{\partial t}$  of the 1st and the  $\frac{\partial}{\partial x}$  of the 2nd of the system (10) and adding them together). But again, we will *not* exclude the possibility of the two systems passing through *different* electric fields in other regions of spacetime (that do *not* contain the observation point), i.e. for  $(x', t') \neq (x, t)$ . In fact, this possibility **will come out naturally** from a careful solution of the basic system (10); it is for example straightforward for the reader to immediately verify that the results (11) or (12) that will be derived below (and will contain contributions of electric field-differences from remote regions of spacetime) indeed satisfy the basic input system of equations (10), something that will be explicitly verified below].

Returning to the question on the appropriate  $\Lambda$  that takes us from the set  $(A_1, \phi_1)$  to the set  $(A_2, \phi_2)$ , we note again that, in cases of static vector potentials ( $A(x)$ 's) *and* spatially uniform scalar potentials ( $\phi(t)$ 's) the form usually given for  $\Lambda$  is the well-known

$$\Lambda(x, t) = \Lambda(x_0, t_0) + \int_{x_0}^x A(x') dx' - c \int_{t_0}^t \phi(t') dt' \quad (8)$$

with  $A(x) = A_2(x) - A_1(x)$  and  $\phi(t) = \phi_2(t) - \phi_1(t)$  (and it can be viewed as a combination of (3) and (4), being immediately applicable to the description of cases of *combined* magnetic and electric Aharonov-Bohm effects).

But as already noted, even in the most general case, with  $t$ -dependent  $A$ 's and  $x$ -dependent  $\phi$ 's (and with the variables  $x$  and  $t$  being **completely uncorrelated**), it is often stated in the literature that the appropriate  $\Lambda$  has a form that is a plausible extension of (8), namely

$$\Lambda(x, t) = \Lambda(x_0, t_0) + \int_{x_0}^x [A_2(x', t) - A_1(x', t)] dx' - c \int_{t_0}^t [\phi_2(x, t') - \phi_1(x, t')] dt', \quad (9)$$

[with eq. (57) of Ref.[4], taken for  $\mathbf{v} = 0$ , being a very good example to point to, since that article does not use a path-integral language, but a canonical formulation with uncorrelated variables]. And as already pointed out in Section II, this form is certainly *incorrect* for uncorrelated variables  $x$  and  $t$  (the reader can easily verify that the system of equations (10) below is *not* satisfied by (9) – see again Section II, especially the paragraph after eq.(6)). We will find in the present work that the correct form consists of two major modifications of (9):

(i) The first leads to the natural appearance of a *path* that continuously connects initial and final points in spacetime, a property that (9) *does not have* [indeed, if the integration curves of (9) are drawn in the  $(x, t)$ -plane, they do *not* form a continuous path from  $(x_0, t_0)$  to  $(x, t)$ ].

(ii) And the second modification is highly nontrivial: it consists of nonlocal contributions of classical electric field-differences from remote regions of spacetime. We will discuss below the consequences of these terms and we will later show that such nonlocal contributions also appear (in an extended form) in more general situations, i.e. they are also present in higher spatial dimensionality (and they then also involve remote magnetic fields in combination with the electric ones); these lead to modifications of ordinary Aharonov-Bohm behaviors or have other important consequences, one of them being a natural remedy of Causality “paradoxes” in time-dependent Aharonov-Bohm experiments.

The form (9) commonly used is of course motivated by the well-known Wu & Yang[8] nonintegrable phase factor, that has a phase equal to  $\int A_\mu dx^\mu = \int A dx - c \int \phi dt$ , a form that appears naturally within the framework of path-integral treatments, or generally in physical situations where narrow wavepackets are implicitly assumed for the quantum particle: the integrals appearing in (9) are then taken along particle trajectories (hence spatial and temporal variables *not* being uncorrelated, but being connected in a particular manner  $x(t)$  to produce the path; all integrals are therefore basically only time-integrals). But even then, eq.(9) is valid only when these trajectories are always (in time) and everywhere (in space) inside identical classical fields for the two (mapped) systems. Here, however, we will be focusing on what a canonical (and not a path-integral or other semiclassical) treatment leads to; this will cover the general case of arbitrary wavefunctions that can even be completely spread-out in space, and will also allow the particle to travel through different electric fields for the two systems in remote spacetime regions (e.g.  $E_2(x, t') \neq E_1(x, t)$  if  $t' < t$  etc.).

It is therefore clear that finding the appropriate  $\Lambda(x, t)$  that achieves the above mapping *in full generality* will require a careful solution of the system of PDEs (2), applied to only one spatial variable, namely

$$\frac{\partial \Lambda(x, t)}{\partial x} = A(x, t) \quad \text{and} \quad -\frac{1}{c} \frac{\partial \Lambda(x, t)}{\partial t} = \phi(x, t) \quad (10)$$

(with  $A(x, t) = A_2(x, t) - A_1(x, t)$  and  $\phi(x, t) = \phi_2(x, t) - \phi_1(x, t)$ ). This system is

underdetermined in the sense that we only have knowledge of  $\Lambda$  at an initial point  $(x_0, t_0)$  and with no further boundary conditions (hence multiplicities of solutions being generally expected, see below). By following a careful procedure of integrations[5] we finally obtain 2 distinct solutions (depending on which equation we integrate first): the first solution is

$$\Lambda(x, t) = \Lambda(x_0, t_0) + \int_{x_0}^x A(x', t) dx' - c \int_{t_0}^t \phi(x_0, t') dt' + \left\{ c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t') + g(x) \right\} + \tau(t_0) \quad (11)$$

with  $g(x)$  required to be chosen so that the quantity  $\left\{ c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t') + g(x) \right\}$  is independent of  $x$ , and (from an inverted route of integrations) the second solution turns out to be

$$\Lambda(x, t) = \Lambda(x_0, t_0) + \int_{x_0}^x A(x', t_0) dx' - c \int_{t_0}^t \phi(x, t') dt' + \left\{ -c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t') + \hat{g}(t) \right\} + \chi(x_0) \quad (12)$$

with  $\hat{g}(t)$  to be chosen in such a way that  $\left\{ -c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t') + \hat{g}(t) \right\}$  is independent of  $t$ .

In the above  $E = (E_2 - E_1)$  is the difference of electric fields in the two systems, which can be nonvanishing at regions remote to the observation point  $(x, t)$  (see examples later below). (Note again that at the point of observation  $E(x, t) = 0$ , signifying the basic fact that the fields in the two systems are identical at the observation point  $(x, t)$ ). The constant last terms in both solutions can be shown to be related to possible multiplicities of  $\Lambda$  (for a full discussion see [5]) and they are zero in simple-connected spacetimes. Also note again that the integrations of potentials in (11) and (12) indeed form paths that continuously connect  $(x_0, t_0)$  to  $(x, t)$  in the  $xt$ -plane (the red-arrow and green-arrow paths of Fig.1(a)), a property that the incorrectly used solution (9) does *not* have.

By “going backwards” one can directly verify that (11) or (12) are indeed solutions of the basic system of PDEs (10), **even for any nonzero**  $E(x', t')$  (in regions  $(x', t') \neq (x, t)$ ). Indeed, if we call our first solution (eq.(11)) for simple-connected spacetime  $\Lambda_1$ , namely

$$\Lambda_1(x, t) = \Lambda_1(x_0, t_0) + \int_{x_0}^x A(x', t) dx' - c \int_{t_0}^t \phi(x_0, t') dt' + \left\{ c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t') + g(x) \right\} \quad (13)$$

with  $g(x)$  chosen so that  $\left\{ c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t') + g(x) \right\}$  is independent of  $x$ , then we have (even for  $E(x', t') \neq 0$  for  $(x', t') \neq (x, t)$ ):

**A)**  $\frac{\partial \Lambda_1(x, t)}{\partial x} = A(x, t)$  satisfied trivially ✓  
(because the quantity  $\{...\}$  is independent of  $x$ ).

**B)**  $-\frac{1}{c} \frac{\partial \Lambda_1(x, t)}{\partial t} = -\frac{1}{c} \int_{x_0}^x \frac{\partial A(x', t)}{\partial t} dx' + \phi(x_0, t) - \int_{x_0}^x E(x', t) dx' - \frac{1}{c} \frac{\partial g(x)}{\partial t},$

(the last term being trivially zero,  $\frac{\partial g(x)}{\partial t} = 0$ ), and then with the substitution

$$-\frac{1}{c} \frac{\partial A(x', t)}{\partial t} = \frac{\partial \phi(x', t)}{\partial x'} + E(x', t)$$

we obtain

$$-\frac{1}{c} \frac{\partial \Lambda_1(x, t)}{\partial t} = \int_{x_0}^x \frac{\partial \phi(x', t)}{\partial x'} dx' + \int_{x_0}^x E(x', t) dx' + \phi(x_0, t) - \int_{x_0}^x E(x', t) dx'.$$

(i) We see that the 2nd and 4th terms of the rhs *cancel each other*, and

(ii) the 1st term of the rhs is  $\int_{x_0}^x \frac{\partial \phi(x', t)}{\partial x'} dx' = \phi(x, t) - \phi(x_0, t).$

Hence finally

$$-\frac{1}{c} \frac{\partial \Lambda_1(x, t)}{\partial t} = \phi(x, t). \quad \checkmark$$

We have directly shown therefore that the basic system of PDEs (10) is indeed satisfied by our **generalized** solution  $\Lambda_1(x, t)$ , **even for any nonzero**  $E(x', t')$  (in regions  $(x', t') \neq (x, t)$ ). (Once again note, however, that at the point of observation  $E(x, t) = 0$ , indicating the essential fact that the fields in the two systems are equal (recall that  $E = E_2 - E_1$ ) at the observation point  $(x, t)$ . It should be noted that the function  $g(x)$  owes its existence to the fact that the spacetime point of observation  $(x, t)$  is outside the  $E$ -distribution (hence the term *nonlocal*, used for the effect of the field-difference  $E$  on the phases), and the reader can clearly see this in the “striped”  $E$ -distributions of the examples that follow later in this Section.

In a completely analogous way, one can easily see that our alternative solution (eq.(12)) also satisfies the basic system of PDEs above. Indeed, if we call our second (alternative) solution (eq.(12)) for simple-connected spacetime  $\Lambda_2$ , namely

$$\Lambda_2(x, t) = \Lambda_2(x_0, t_0) + \int_{x_0}^x A(x', t_0) dx' - c \int_{t_0}^t \phi(x, t') dt' + \left\{ -c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t') + \hat{g}(t) \right\} \quad (14)$$

with  $\hat{g}(t)$  chosen so that  $\left\{ -c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t') + \hat{g}(t) \right\}$  is independent of  $t$ , then we have (even for  $E(x', t') \neq 0$  for  $(x', t') \neq (x, t)$ ):

**A)**  $-\frac{1}{c} \frac{\partial \Lambda_2(x, t)}{\partial t} = \phi(x, t)$  satisfied trivially  $\checkmark$   
(because the quantity  $\{....\}$  is independent of  $t$ ).

**B)**  $\frac{\partial \Lambda_2(x, t)}{\partial x} = A(x, t_0) - c \int_{t_0}^t \frac{\partial \phi(x, t')}{\partial x} dt' - c \int_{t_0}^t E(x, t') dt' + \frac{\partial \hat{g}(t)}{\partial x}$ ,

(the last term being trivially zero,  $\frac{\partial \hat{g}(t)}{\partial x} = 0$ ), and then with the substitution

$$\frac{\partial \phi(x, t')}{\partial x} = -E(x, t') - \frac{1}{c} \frac{\partial A(x, t')}{\partial t'}$$

we obtain

$$\frac{\partial \Lambda_2(x, t)}{\partial x} = A(x, t_0) + c \int_{t_0}^t E(x, t') dt' + \int_{t_0}^t \frac{\partial A(x, t')}{\partial t'} dt' - c \int_{t_0}^t E(x, t') dt'.$$

(i) We see that the 2nd and 4th terms of the rhs *cancel each other*, and

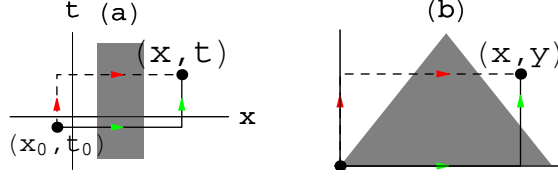
(ii) the 3rd term of the rhs is  $\int_{t_0}^t \frac{\partial A(x, t')}{\partial t'} dt' = A(x, t) - A(x, t_0)$ .

Hence finally

$$\frac{\partial \Lambda_2(x, t)}{\partial x} = A(x, t). \quad \checkmark$$

Once again, all the above are true for any nonzero  $E(x', t')$  (in regions  $(x', t') \neq (x, t)$ ) for arbitrary analytical dependence of the remote field-difference on its arguments.

Let us now note that in (11) and (12) the placement of  $x_0$  and  $t_0$  gives a “path-sense” to the line integrals in each solution (each path consisting of 2 perpendicular line segments connecting  $(x_0, t_0)$  to  $(x, t)$ , with solution (11) having a clockwise and solution (12) a



counter-clockwise sense, see red and green arrow paths in Fig.1); this way a natural *rectangle* is formed, within which the enclosed “electric fluxes” in spacetime appear to be crucial (showing up as nonlocal contributions of the electric field-differences from regions  $(x', t')$  of space and time *that are remote to the observation point*  $(x, t)$ ). These nonlocal terms in  $\Lambda$  have a direct effect on the wavefunction-phases at  $(x, t)$ . The actual manner in which this happens is determined by the functions  $g(x)$  or  $\hat{g}(t)$ —these must be chosen in such a way as to satisfy their respective conditions in (11) and (12). To see how these functions ( $g(x)$  or  $\hat{g}(t)$ ) are actually determined, and what form the above solutions take in nontrivial cases (and how they give new results, i.e. *not differing from the usual ones by a mere constant*) let us first take examples of striped  $E$ -distributions in spacetime:

**(a)** For the case of the extended *vertical* strip (parallel to the  $t$ -axis) of Fig.1(a) (the case of a one-dimensional capacitor that is (arbitrarily and variably) charged for all time), then, for  $x$  located outside (and on the right of) the capacitor, the quantity  $c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t')$  in  $\Lambda_1$  is *already independent of  $x$*  (since a displacement of the  $(x, t)$ -corner of the rectangle to the right, along the  $x$ -direction, does not change the enclosed “electric flux”, see Fig.1(a)); hence in this case the function  $g(x)$  can be taken as  $g(x) = 0$  (up to a constant  $C$ ), because then the condition for  $g(x)$  stated in the solution (11) (namely, that the quantity in brackets must be independent of  $x$ ) is indeed satisfied. (Note again that the above  $x$ -independence of the enclosed “electric flux” is important for the existence of  $g(x)$ ).

So for this setup, the nonlocal term in solution  $\Lambda_1$  *survives* (the quantity in brackets is nonvanishing), but *it is not constant*: this enclosed flux depends on  $t$  (since the enclosed flux *does change* with a displacement of the  $(x, t)$ -corner of the rectangle upwards, along the  $t$ -direction). Hence, by looking at the alternative solution  $\Lambda_2(x, t)$ , the quantity  $c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t')$  is dependent on  $t$ , so that  $\hat{g}(t)$  must be chosen as  $\hat{g}(t) =$

$+c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t')$  (up to the same constant  $C$ ) in order to *cancel* this  $t$ -dependence, so that its own condition stated in the solution (12) (namely, that the quantity in brackets must be independent of  $t$ ) is indeed satisfied; as a result, the quantity in brackets in solution  $\Lambda_2$  disappears and there is no nonlocal contribution in  $\Lambda_2$  (for  $C = 0$ ). (If we had used a  $C \neq 0$ , the nonlocal contributions would be differently shared between the two solutions, but without changing the Physics when we take the *difference* of the two solutions, see below).

With these choices of  $\hat{g}(t)$  and  $g(x)$ , we already have new results (compared to the standard ones of the integrals of potentials). I.e. one of the two solutions, namely  $\Lambda_1$  **is** affected nonlocally by the enclosed flux (and this flux is **not** constant). Spelled out clearly, the two results are:

$$\Lambda_1(x, t) = \Lambda_1(x_0, t_0) + \int_{x_0}^x A(x', t) dx' - c \int_{t_0}^t \phi(x_0, t') dt' + c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t') + C \quad (15)$$

$$\Lambda_2(x, t) = \Lambda_2(x_0, t_0) + \int_{x_0}^x A(x', t_0) dx' - c \int_{t_0}^t \phi(x, t') dt' + C. \quad (16)$$

And now it is easy to note that, if we subtract the two solutions  $\Lambda_1$  and  $\Lambda_2$  (and of course assume, as usual, single-valuedness of  $\Lambda$  at the initial point  $(x_0, t_0)$ , i.e.  $\Lambda_1(x_0, t_0) = \Lambda_2(x_0, t_0)$ ) the result is *zero* (i.e. it so happens that the electric flux determined by the potential-integrals is exactly cancelled by the nonlocal term of electric fields (i.e. the term that survives in  $\Lambda_1$  above)), a cancellation effect that is important and that will be generalized in later Sections.

(b) In the “dual case” of an extended *horizontal* strip - parallel to the  $x$ -axis (that corresponds to a nonzero electric field in all space that has however a finite duration  $T$ ), the proper choices (for observation time instant  $t > T$ ) are basically reverse (i.e. we can now take  $\hat{g}(t) = 0$  and  $g(x) = -c \int_{t_0}^t dt' \int_{x_0}^x dx' E(x', t')$  (since the “electric flux” enclosed in the “observation rectangle” now depends on  $x$ , but not on  $t$ ), with both choices always up to a common constant) and once again we can easily see, upon subtraction of the two solutions, a similar cancellation effect. In this case again, the results are also new (a nonlocal term survives now in  $\Lambda_2$ ). Again spelled out clearly, these are:



$$\Lambda_1(x, t) = \Lambda_1(x_0, t_0) + \int_{x_0}^x A(x', t) dx' - c \int_{t_0}^t \phi(x_0, t') dt' + C \quad (17)$$

$$\Lambda_2(x, t) = \Lambda_2(x_0, t_0) + \int_{x_0}^x A(x', t_0) dx' - c \int_{t_0}^t \phi(x, t') dt' - c \int_{x_0}^x dx' \int_{t_0}^t dt' E(x', t') + C \quad (18)$$

their difference also being zero.

(c) And if we want cases that are more involved (with the nonlocal contributions appearing nontrivially in **both** solutions  $\Lambda_1$  and  $\Lambda_2$  and with  $g(x)$  and  $\hat{g}(t)$  not being “immediately visible”) we must again consider different shapes of  $E$ -distribution. One such case (a triangular  $E$ -distribution) is shown in Fig.1(b) (for the corresponding magnetic case to be discussed in the next Section, which however is completely analogous); in this case the enclosed flux depends on *both*  $x$  and  $t$  (but can be shown to be separable, so that the functions  $g(x)$  and  $\hat{g}(t)$  still exist and can easily be found in closed analytical form, see next Section). As for the last constant terms  $\tau(t_0)$  and  $\chi(x_0)$  (what we will call “multiplicities”), these are only present (nonvanishing) when  $\Lambda$  is expected to be multivalued, i.e. in cases of motion in multiple-connected spacetimes, and are then related to the fluxes in the inaccessible regions: in the electric Aharonov-Bohm setup, the prototype of multiple-connectivity in spacetime[9], it turns out[5] that  $\tau(t_0) = -\chi(x_0) =$  enclosed (and here inaccessible) “electric flux”, and if these values are substituted in (11) and (12) they cancel out the new nonlocal terms and lead to the usual electric Aharonov-Bohm result (of mere integrals over potentials). As for other, more esoteric properties of the new solutions in simple-connected spacetimes, it can be rigorously shown[5] that solutions (11) and (12) are actually *equal*, because  $g(x)$  turns out to be equal to the  $t$ -independent bracket of (12), and  $\hat{g}(t)$  turns out to be equal to the  $x$ -independent bracket of (11), the nonlocal terms having therefore the tendency to exactly cancel the “Aharonov-Bohm terms” (this being true *for arbitrary shapes and analytical form of*  $E(x, t)$ ).

#### IV. 2-D STATIC CASE

After having discussed fully the simple  $(x, t)$ -case, let us for completeness give the analogous (Euclidian-rotated in 4-D spacetime) derivation for  $(x, y)$ -variables and briefly discuss

the properties of the simpler static solutions in spatial two-dimensionality. We will simply need to apply the same methodology (of solution of a system of PDEs) to such static spatially two-dimensional cases (so that now different (remote) magnetic fields for the two systems, perpendicular to the 2-D space, will arise). For such cases we need to solve the system of PDEs that is now of the form

$$\frac{\partial \Lambda(x, y)}{\partial x} = A_x(x, y) \quad \text{and} \quad \frac{\partial \Lambda(x, y)}{\partial y} = A_y(x, y). \quad (19)$$

By following then a similar procedure of integrations[5] we obtain the following general solution

$$\Lambda(x, y) = \Lambda(x_0, y_0) + \int_{x_0}^x A_x(x', y) dx' + \int_{y_0}^y A_y(x_0, y') dy' + \left\{ \int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y') + g(x) \right\} + f(y_0) \quad (20)$$

with  $g(x)$  chosen so that  $\left\{ \int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y') + g(x) \right\}$  : is independent of  $x$ ,

a result that applies to cases where the particle goes through *different* perpendicular magnetic fields  $\mathbf{B}_2$  and  $\mathbf{B}_1$  in spatial regions that are remote to (i.e. do not contain) the observation point  $(x, y)$  (and in the above  $B_z = (\mathbf{B}_2 - \mathbf{B}_1)_z$ ). The reader should note that the first 3 terms of (20) are the (total) Dirac phase along two perpendicular segments that continuously connect the initial point  $(x_0, y_0)$  to the point of observation  $(x, y)$ , in a clockwise sense (see for example the red-arrow paths in Fig.1(b)). But apart from this Dirac phase, we also have nonlocal contributions from  $B_z$  and its flux within the “observation rectangle” (see i.e. the rectangle being formed by the red- and green-arrow paths in Fig.1(b)). Below we will directly verify that (20) is indeed a solution of (19) (even for  $B_z(x', y') \neq 0$  for  $(x', y') \neq (x, y)$ ). Alternatively, by following the reverse route of integrations, we finally obtain the following alternative general solution

$$\Lambda(x, y) = \Lambda(x_0, y_0) + \int_{x_0}^x A_x(x', y_0) dx' + \int_{y_0}^y A_y(x, y') dy' + \left\{ - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y') + h(y) \right\} + \hat{h}(x_0) \quad (21)$$

with  $h(y)$  chosen so that  $\left\{ -\int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y') + h(y) \right\}$  : is independent of  $y$ ,

and again the reader should note that, apart from the first 3 terms (the (total) Dirac phase along the two other (alternative) perpendicular segments (connecting  $(x_0, y_0)$  to  $(x, y)$ ), now *in a counterclockwise sense* (the green-arrow paths in Fig.1(b))), we also have nonlocal contributions from the flux of  $B_z$  that is enclosed within the same “observation rectangle” (that is naturally defined by the four segments of the two solutions (Fig.1(b))).

In all the above,  $A_x$  and  $A_y$  are the Cartesian components of  $\mathbf{A}(\mathbf{r}) = \mathbf{A}(x, y) = \mathbf{A}_2(\mathbf{r}) - \mathbf{A}_1(\mathbf{r})$ , and, as already mentioned,  $B_z$  is the difference between (perpendicular) magnetic fields that the two systems may experience in regions that *do not contain* the observation point  $(x, y)$  (i.e.  $B_z(x', y') = (\mathbf{B}_2(x', y') - \mathbf{B}_1(x', y'))_z = \frac{\partial A_y(x', y')}{\partial x'} - \frac{\partial A_x(x', y')}{\partial y'}$ , and, although at the point of observation  $(x, y)$  we have  $B_z(x, y) = 0$  (already emphasized in the Introductory Sections), this  $B_z(x', y')$  can be nonzero for  $(x', y') \neq (x, y)$ ). It should be noted that it is because of  $B_z(x, y) = 0$  that the functions  $g(x)$  and  $h(y)$  of (20) and (21) can be found, and the new solutions therefore exist (and are nontrivial). For the impatient reader, simple physical examples with the associated analytical forms of  $g(x)$  and  $h(y)$  (derived in detail) are given later in this Section.

One can again show that the 2 solutions are equal for simple-connected space (when the last constant terms  $f(y_0)$  and  $\hat{h}(x_0)$  are vanishing), and for multiple-connectivity the values of the multiplicities  $f(y_0)$  and  $\hat{h}(x_0)$  cancel out the nonlocalities and reduce the above to the usual result of mere  $A$ -integrals along the 2 paths (i.e. two simple Dirac phases).

A direct “backwards” verification that (20) and (21) do indeed satisfy the basic system (19) (**even for cases with  $B_z \neq 0$**  in remote regions of space) can be made along similar lines to the ones of the last Section: for simple-connected space, let us call our solution (20)  $\Lambda_3$ , namely

$$\Lambda_3(x, y) = \Lambda_3(x_0, y_0) + \int_{x_0}^x A_x(x', y) dx' + \int_{y_0}^y A_y(x_0, y') dy' + \left\{ \int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y') + g(x) \right\} \quad (22)$$

with  $g(x)$  chosen so that  $\left\{ \int_{y_0}^y \int_{x_0}^x B_z + g(x) \right\}$  is independent of  $x$ . We then have (even for

$B_z(x', y') \neq 0$  for  $(x', y') \neq (x, y)$ ):

**A)**  $\frac{\partial \Lambda_3(x, y)}{\partial x} = A_x(x, y)$  satisfied trivially ✓

(because the quantity  $\{\dots\}$  is independent of  $x$ ).

**B)**  $\frac{\partial \Lambda_3(x, y)}{\partial y} = \int_{x_0}^x \frac{\partial A_x(x', y)}{\partial y} dx' + A_y(x_0, y) + \int_{x_0}^x B_z(x', y) dx' + \frac{\partial g(x)}{\partial y},$

(the last term being trivially zero,  $\frac{\partial g(x)}{\partial y} = 0$ ), and then with the substitution

$$\frac{\partial A_x(x', y)}{\partial y} = \frac{\partial A_y(x', y)}{\partial x'} - B_z(x', y)$$

we obtain

$$\frac{\partial \Lambda_3(x, y)}{\partial y} = \int_{x_0}^x \frac{\partial A_y(x', y)}{\partial x'} dx' - \int_{x_0}^x B_z(x', y) dx' + A_y(x_0, y) + \int_{x_0}^x B_z(x', y) dx'.$$

(i) We see that the 2nd and 4th terms of the right-hand-side (rhs) *cancel each other*, and

(ii) the 1st term of the rhs is  $\int_{x_0}^x \frac{\partial A_y(x', y)}{\partial x'} dx' = A_y(x, y) - A_y(x_0, y).$

Hence finally

$$\frac{\partial \Lambda_3(x, y)}{\partial y} = A_y(x, y). \quad \checkmark$$

We have directly shown therefore (by “going backwards”) that the basic system of PDEs (19) is indeed satisfied by our generalized solution  $\Lambda_3(x, y)$ , even for any nonzero  $B_z(x', y')$  (in regions  $(x', y') \neq (x, y)$ ; recall that always  $B_z(x, y) = 0$ ). To fully appreciate the above simple proof, the reader is again urged to look at the cases of “striped”  $B_z$ -distributions later below, the point of observation  $(x, y)$  always lying outside the strips, so that the above function  $g(x)$  can easily be determined, and the new solutions really *exist* - and they are nontrivial.

In a completely analogous way, one can easily see that our alternative solution (eq.(21)) also satisfies the basic system of PDEs above. Indeed, if we call this second static solution  $\Lambda_4$ , namely

$$\Lambda_4(x, y) = \Lambda_4(x_0, y_0) + \int_{x_0}^x A_x(x', y_0) dx' + \int_{y_0}^y A_y(x, y') dy' + \left\{ - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y') + h(y) \right\} \quad (23)$$

with  $h(y)$  chosen so that  $\left\{ - \int_{x_0}^x \int_{y_0}^y B_z + h(y) \right\}$  : is independent of  $y$ , then we have (even for  $B_z(x', y') \neq 0$  for  $(x', y') \neq (x, y)$ ):

A)  $\frac{\partial \Lambda_4(x, y)}{\partial y} = A_y(x, y)$  satisfied trivially ✓

(because the quantity  $\{...\}$  is independent of  $y$ ).

B)  $\frac{\partial \Lambda_4(x, y)}{\partial x} = A_x(x, y_0) + \int_{y_0}^y \frac{\partial A_y(x, y')}{\partial x} dy' - \int_{y_0}^y B_z(x, y') dy' + \frac{\partial h(y)}{\partial x},$

(the last term being trivially zero,  $\frac{\partial h(y)}{\partial x} = 0$ ), and then with the substitution

$$\frac{\partial A_y(x, y')}{\partial x} = \frac{\partial A_x(x, y')}{\partial y'} + B_z(x, y')$$

we obtain

$$\frac{\partial \Lambda_4(x, y)}{\partial x} = A_x(x, y_0) + \int_{y_0}^y \frac{\partial A_x(x, y')}{\partial y'} dy' + \int_{y_0}^y B_z(x, y') dy' - \int_{y_0}^y B_z(x, y') dy'.$$

(i) We see that the last two terms of the rhs *cancel each other*, and

(ii) the 2nd term of the rhs is  $\int_{y_0}^y \frac{\partial A_x(x, y')}{\partial y'} dy' = A_x(x, y) - A_x(x, y_0).$

Hence finally

$$\frac{\partial \Lambda_4(x, y)}{\partial x} = A_x(x, y). \quad \checkmark$$

Once again, all the above are true for any nonzero  $B_z(x', y')$  (in regions  $(x', y') \neq (x, y)$ ) for arbitrary analytical dependence of the remote field-difference on its arguments. And for a clearer understanding of this proof let us now turn to the “striped” examples promised earlier.

In order to see again how the above solutions appear in nontrivial cases (and how they give completely new results, i.e. *not differing from the usual ones (i.e. from the Dirac phase) by a mere constant*) let us first take examples of striped  $B_z$ -distributions in space:

(a) For the case of an extended *vertical* strip - parallel to the  $y$ -axis, such as in Fig.1(a) (imagine  $t$  replaced by  $y$ ) (i.e. for the case that the particle has actually passed through nonzero  $B_z$ , hence through *different* magnetic fields in the two (mapped) systems), then, for  $x$  located outside (and on the right of) the strip, the quantity  $\int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y')$  in  $\Lambda_3$  is *already independent of  $x$*  (since a displacement of the  $(x, y)$ -corner of the rectangle to the right, along the  $x$ -direction, does not change the enclosed magnetic flux – see Fig. 1(a) for the analogous  $(x, t)$ -case discussed earlier). Indeed, in this case the above quantity (the enclosed flux within the “observation rectangle”) does not depend on the  $x$ -position of the observation point, but on the positioning of the boundaries of the  $B_z$ -distribution in the

$x$ -direction (better, on the constant width of the strip) – as the  $x$ -integral does not give any further contribution when the dummy variable  $x'$  goes out of the strip. In fact, in this case the enclosed flux depends on  $y$  as we discuss below (but, again, not on  $x$ ). Hence, for this case, the function  $g(x)$  can be easily determined: it can be taken as  $g(x) = 0$  (up to a constant  $C$ ), because then the condition for  $g(x)$  stated in solution (20) (namely, that the quantity in brackets must be independent of  $x$ ) is indeed satisfied.

We see therefore above that for this setup, the nonlocal term in solution  $\Lambda_3$  *survives* (the quantity in brackets is nonvanishing), but *it is not constant*: as already noted, this enclosed flux depends on  $y$  (since the enclosed flux *does change* with a displacement of the  $(x, y)$ -corner of the rectangle upwards, along the  $y$ -direction, as the  $y$ -integral *is* affected by the positioning of  $y$  – the higher the positioning of the observation point the more flux is enclosed inside the observation rectangle). Hence, by looking at the alternative solution  $\Lambda_4(x, y)$ , the quantity  $\int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y')$  is *dependent on  $y$* , so that  $h(y)$  must be chosen as  $h(y) = + \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y')$  (up to the same constant  $C$ ) in order to *cancel* this  $y$ -dependence, so that its own condition stated in solution (21) (namely, that the quantity in brackets must be independent of  $y$ ) is indeed satisfied; as a result, the quantity in brackets in solution  $\Lambda_4$  disappears and there is no nonlocal contribution in  $\Lambda_4$  (for  $C = 0$ ). (If we had used a  $C \neq 0$ , the nonlocal contributions would be shared between the two solutions in a different manner, but without changing the Physics when we take the *difference* of the two solutions (see below)). [The crucial point in the above for the existence of  $g(x)$  and  $h(y)$  is, once again, the fact that  $B_z = 0$  at  $(x, y)$ , combined with the sharp boundaries of the nonvanishing  $B_z$ -region].

With these choices of  $h(y)$  and  $g(x)$ , we already have new results (compared to the standard ones of the integrals of potentials). I.e. one of the two solutions, namely  $\Lambda_3$  **is** affected nonlocally by the enclosed flux (and this flux is **not** constant). Spelled out clearly, the two results are:

$$\Lambda_3(x, y) = \Lambda_3(x_0, y_0) + \int_{x_0}^x A_x(x', y) dx' + \int_{y_0}^y A_y(x_0, y') dy' + \int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y') + C \quad (24)$$

$$\Lambda_4(x, y) = \Lambda_4(x_0, y_0) + \int_{x_0}^x A_x(x', y_0) dx' + \int_{y_0}^y A_y(x, y') dy' + C. \quad (25)$$

And now it is easy to note that, if we subtract the two solutions  $\Lambda_3$  and  $\Lambda_4$ , the result is *zero* (because the line integrals of the vector potential  $\mathbf{A}$  in the two solutions are in opposite senses in the  $(x, y)$  plane, hence their difference leads to a *closed* line integral of  $\mathbf{A}$ , which is in turn equal to the enclosed magnetic flux, and this flux always happens to be of opposite sign from that of the enclosed flux that explicitly appears as a nonlocal contribution of the  $B_z$ -fields (i.e. the term that survives in  $\Lambda_3$  above). Hence, the two solutions are *equal*. [We of course everywhere assumed, as usual, single-valuedness of  $\Lambda$  at the initial point  $(x_0, y_0)$ , i.e.  $\Lambda_1(x_0, y_0) = \Lambda_2(x_0, y_0)$ ; matters of multivaluedness of  $\Lambda$  at the observation point  $(x, y)$  will be addressed later below].

It is interesting that, formally speaking, the above *equality* of the two solutions is due to the fact that the  $x$ -independent quantity in brackets of the 1st solution (20) is equal to the function  $h(y)$  of the 2nd solution (21), and the  $y$ -independent quantity in brackets of the 2nd solution (21) is equal to the function  $g(x)$  of the first solution (20). This turns out to be a general behavioral pattern of the two solutions in simple-connected space, that is valid for any shape (and any analytical form) of  $B_z$ -distribution. But most importantly, it should be noted that this vanishing of  $\Lambda_3(x, y) - \Lambda_4(x, y)$  is a cancellation effect that is emphasized further later below and discussed in completely *physical terms*.

(b) In the “dual case” of an extended *horizontal* strip - parallel to the  $x$ -axis, the proper choices (for  $y$  above the strip) are basically reverse (i.e. we can now take  $h(y) = 0$  and  $g(x) = -\int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y')$  (since the flux enclosed in the rectangle now depends on  $x$ , but not on  $y$ ), with both choices always up to a common constant) and once again we can easily see, upon subtraction of the two solutions, a similar cancellation effect. In this case as well, the results are again new (a nonlocal term survives now in  $\Lambda_4$ ). Again spelled out clearly, these are:

$$\Lambda_3(x, y) = \Lambda_3(x_0, y_0) + \int_{x_0}^x A_x(x', y) dx' + \int_{y_0}^y A_y(x_0, y') dy' + C \quad (26)$$

$$\Lambda_4(x, y) = \Lambda_4(x_0, y_0) + \int_{x_0}^x A_x(x', y_0) dx' + \int_{y_0}^y A_y(x, y') dy' - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y') + C \quad (27)$$

(their difference also being zero). Again here the crucial point is that, because the  $B_z$ -configuration does *not contain the point*  $(x, y)$ , a displacement of this observation point upwards does *not* change the flux inside the “observation rectangle”; this makes the new solutions (i.e the functions  $g(x)$  and  $h(y)$ ) exist.

(c) And if we want cases that are more involved (i.e. with the nonlocal contributions appearing nontrivially in **both** solutions  $\Lambda_3$  and  $\Lambda_4$  and with  $g(x)$  and  $h(y)$  not being “immediately visible”), we must consider different shapes of  $B_z$ -distributions. One such case is a triangular one that is shown in Fig.1(b) (for simplicity an equilateral triangle, and with the initial point  $(x_0, y_0) = (0, 0)$ ) and with the point of observation  $(x, y)$  being fairly close to the triangle’s right side as in the figure. Note that for such a configuration, the part of the magnetic flux that is inside the “observation rectangle” (defined by the right upper corner  $(x, y)$ ) depends on **both**  $x$  **and**  $y$ . It turns out, however, that this  $(x$  and  $y)$ –dependent enclosed flux can be written as a sum of separate  $x$ - and  $y$ -contributions, so that appropriate  $g(x)$  and  $h(y)$  can still be found (each one of them must be chosen so that it only cancels the corresponding variable’s dependence of the enclosed flux). For a homogeneous  $B_z$  it is a rather straightforward exercise to determine this enclosed part, i.e. the common area between the observation rectangle and the equilateral triangle, and from this we can find the appropriate  $g(x)$  that will cancel the  $x$ -dependence, and the appropriate  $h(y)$  that will cancel the  $y$ -dependence. These appropriate choices turn out to be

$$g(x) = B_z \left[ -(\sqrt{3}ax - \frac{\sqrt{3}}{2}x^2) + \frac{\sqrt{3}}{4}a^2 \right] + C \quad (28)$$

and

$$h(y) = B_z \left[ (ay - \frac{y^2}{\sqrt{3}}) - \frac{\sqrt{3}}{4}a^2 \right] + C \quad (29)$$

with  $a$  being the side of the equilateral triangle. We should emphasize that expressions (28) and (29), if combined with (20) or (21), give the nontrivial nonlocal contributions of the field-difference  $B_z$  of the remote magnetic fields on  $\Lambda$  of each solution (hence on the phase of the



wavefunction of each wavepacket travelling along each path) at the observation point  $(x, y)$ , that always lies outside the  $B_z$ -triangle. (We mention again that in the case of completely spread-out states, the equality of the two solutions at the observation point essentially demonstrates the uniqueness (single-valuedness) of the phase in simple-connected space). Further physical discussion of the above cancellations, and a semiclassical interpretation, is given later below.

Finally, in more “difficult” geometries, i.e. when the shape of the  $B_z$ -distribution is such that the enclosed flux does *not* decouple in a sum of separate  $x$ - and  $y$ -contributions, such as cases of circularly shaped  $B_z$ -distributions, it is advantageous to solve the system (19) directly in non-Cartesian (i.e. polar) coordinates[5]. A general comment that can be made for general shapes is that, depending on the geometry of shape of the  $B_z$ -distribution, an appropriate change of variables (to a new coordinate system) may first be needed, so that generalized solutions of the system (19) can be found (namely, so that the enclosed flux inside the *transformed* observation rectangle (i.e. a slice of an annular section in the case of polar coordinates) can be written as a sum of separate (transformed) variables), and then the same methodology (as in the above Cartesian cases) can be followed.

As we saw in the above examples, in case of a striped-distribution of the magnetic field difference  $B_z$ , the functions  $g(x)$  and  $h(y)$  in (20) and (21) have to be chosen in ways that are compatible with their corresponding constraints (stated after (20) and (21)) and are completely analogous to the above discussed  $(x, t)$ -cases. By then taking the *difference* of (20) and (21)) we obtain that the “Aharonov-Bohm phase” (the one originating from the *closed* line integral of  $A$ ’s) is exactly cancelled by the additional nonlocal term of the magnetic fields (that the particle passed through). This is quite reminiscent of the cancellation of phases observed in the early experiments of Werner & Brill[10] for particles passing through a magnetic field (a cancellation between the “Aharonov-Bohm phase” and the semiclassical phase picked up by the trajectories), and our method seems to provide a natural explanation: as our results are general (and for delocalized states in simple-connected space they basically demonstrate the uniqueness of  $\Lambda$ ), they are also valid and applicable to states that describe wavepackets in classical motion, as *was* the case of the Werner & Brill experiments. (A similar cancellation of an electric Aharonov-Bohm phase, that has never been noted in the literature, also occurs for particles passing through a static electric field, and this we will independently prove below, again for semiclassical states). We conclude that, for static



that follows below, where this is proved in detail). Similarly, in the very recent review of Batelaan & Tonomura[13], their Fig.2 shows wavefronts associated to the deflected classical trajectories where it is stated that “the phase shift calculated in terms of the Lorentz force is the same as that predicted by the Aharonov-Bohm effect in terms of the vector potential”. Once more, however, it turns out that the sign of the classical phase-difference is really opposite to the sign of the Aharonov-Bohm phase (see proof below). The phases are not equal as stated by the authors. And it turns out that even “electric analogs” of the above cases also demonstrate this opposite-sign relationship (see proof further below). All the above examples can be viewed as a manifestation of the cancellations that have been found in the present work for *general* (even completely spread-out) quantum states (but in those examples they are just special cases for narrow wavepacket-states in classical motion).

Let us give a brief elementary proof of the above claimed opposite sign-relationships: Indeed, in our Fig.2, the “Aharonov-Bohm phase” due to the flux enclosed between the two classical trajectories (of a particle of charge  $q$ ) is

$$\Delta\varphi^{AB} = 2\pi \frac{q}{e} \frac{\Phi}{\Phi_0}, \quad (30)$$

with  $\Phi_0 = \frac{hc}{e}$  the flux quantum, and  $\Phi \approx BWd$  the enclosed flux between the two trajectories (for small trajectory-deflections), with the deflection originating from the presence of the magnetic strip  $B$  and the associated Lorentz forces. On the other hand, the semiclassical phase difference between the same 2 classical trajectories is  $\Delta\varphi^{semi} = \frac{2\pi}{\lambda} \Delta l$ , with  $\lambda = \frac{h}{mv}$  being the de Broglie wavelength (and  $v$  being the speed of the particle, taken almost constant (as usually done) due to the small deflections), and with  $\Delta l$  being  $\Delta l \approx d \sin \theta \approx d \frac{x_c}{L}$  ( $x_c$  being the (displaced) position of the central fringe on the screen). We have therefore

$$\Delta\varphi^{semi} = \frac{2\pi}{\lambda} d \frac{x_c}{L}. \quad (31)$$

Now, the Lorentz force (exerted only during the passage through the thin magnetic strip, hence only during a time interval  $\Delta t = \frac{W}{v}$ ) has a component parallel to the screen (let us call it  $x$ -component) that is given by

$$F_x = \frac{q}{c} (\mathbf{v} \times \mathbf{B})_x = -\frac{q}{c} v B = -\frac{BWq}{c \frac{W}{v}} = -\frac{BWq}{c \Delta t} \quad (32)$$

which shows that there is a change of kinematic momentum (parallel to the screen) equal to  $-\frac{BWq}{c}$ , or, equivalently, a change of parallel speed

$$\Delta v_x = -\frac{BWq}{mc} \quad (33)$$

which is the speed of the central fringe's motion (i.e. its displacement over time along the screen). Although this has been caused by the presence of the thin deflecting magnetic strip, this displacement is occurring uniformly during a time interval  $t = \frac{L}{v}$ , and this time interval must satisfy

$$\Delta v_x = \frac{x_c}{t} \quad (34)$$

(as, for small displacements, the wavepackets travel most of the time in uniform motion, i.e.  $\Delta t \ll t$ ). We therefore have that the central fringe displacement must be  $x_c = \Delta v_x t = -\frac{BWq}{cm} \frac{L}{v}$ , and noting that  $mv = \frac{h}{\lambda}$ , we finally have

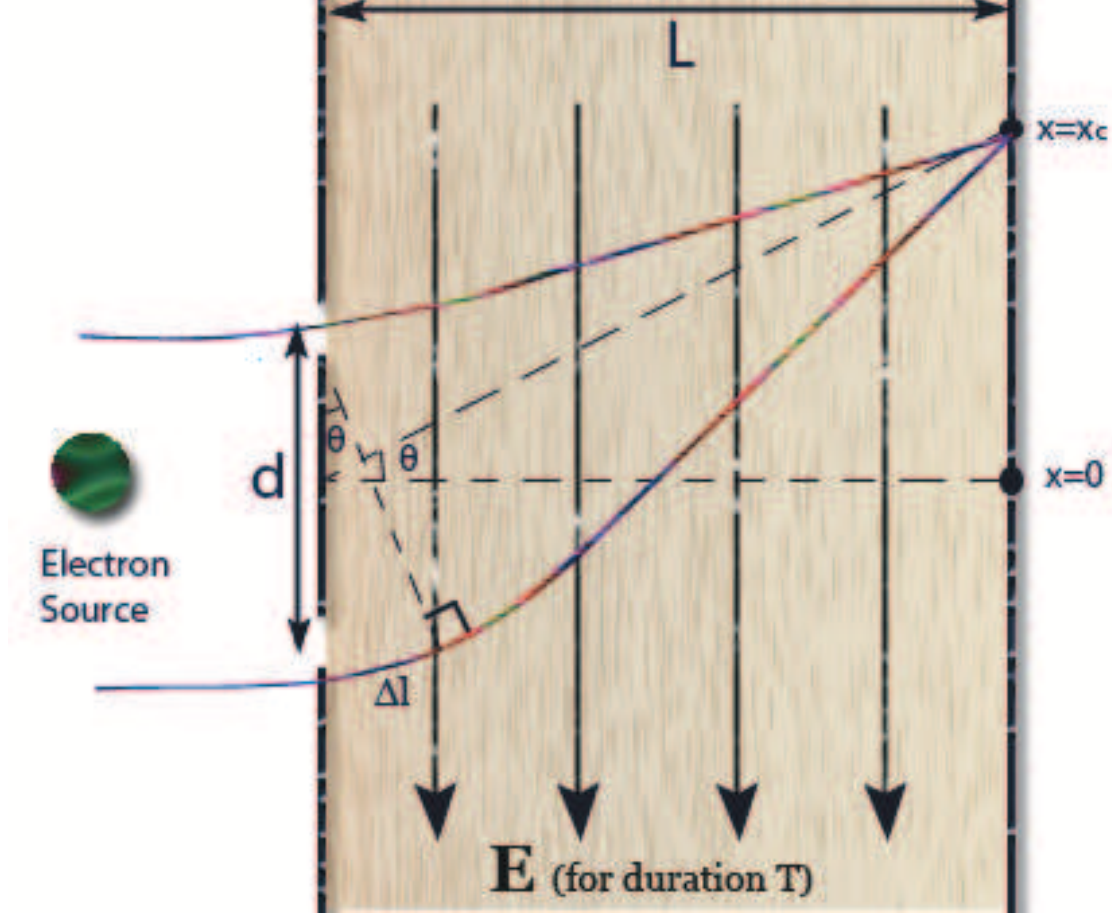
$$x_c = -\frac{BWqL\lambda}{hc}. \quad (35)$$

By substituting (35) into (31), the lengths  $L$  and  $\lambda$  cancel out, and we finally have  $\Delta\varphi^{semi} = -2\pi \frac{q}{e} \frac{BWd}{\frac{hc}{e}}$ , which with  $\frac{hc}{e} = \Phi_0$  the flux quantum, and  $BWd \approx \Phi$  the enclosed flux (always for small trajectory-deflections) gives (through comparison with (30)) our final proof that

$$\Delta\varphi^{semi} = -2\pi \frac{q}{e} \frac{\Phi}{\Phi_0} = -\Delta\varphi^{AB}. \quad (36)$$

[It should be noted that such an opposite sign relation actually tells us that the *total* (semi-classical + AB) phase *difference* is zero at the new position of the central (bright) fringe (after it has been displaced due to the trajectories' deflection), and in this sense the above minus sign should be rather expected.]

The “electric analog” of the above exercise is also outlined below, now with a homogeneous electric field (pointing downwards everywhere in space, but switched on for only a finite duration  $T$ ) on the right of a double-slit apparatus (see our Fig.3): In this case the electric Lorentz force  $qE$  is exerted on the trajectories only during the small time interval  $\Delta t = T$ , which we take to be much shorter ( $T \ll t$ ) than the time of travel  $t = \frac{L}{v}$  (we now have a thin electric strip in *time* rather than the thin magnetic strip in space that we had earlier). The electric type of “Aharonov-Bohm phase” is now



$$\Delta\varphi^{AB} = -2\pi \frac{q}{e} \frac{cT\Delta V}{\Phi_0}, \quad (37)$$

with  $\Delta V$  being the electric potential difference between the two trajectories, hence  $\Delta V \approx Ed$  (again for small trajectory-deflections). On the other hand, the semiclassical phase difference between the two trajectories is again given by (31), but the position  $x_c$  of the central fringe must now be determined by the electric field force  $qE$ : The change of kinematic momentum (always parallel to the screen) is now  $qET$ , hence the analog of (33) is now

$$\Delta v_x = \frac{qET}{m} \quad (38)$$

which if combined with (34) (that is obviously valid in this case as well, again for small deflections, due to the  $\Delta t = T \ll t$ ), and always with  $t = \frac{L}{v}$ , gives that the central fringe displacement must be  $x_c = \Delta v_x t = \frac{qET}{m} \frac{L}{v}$ , and using again  $mv = \frac{h}{\lambda}$ , we finally have the following analog of (35)

$$x_c = \frac{qETL\lambda}{h}. \quad (39)$$

By substituting (39) into (31), the lengths  $L$  and  $\lambda$  again cancel out, and we finally have  $\Delta\varphi^{semi} = 2\pi d \frac{qETL\lambda}{h} = 2\pi \frac{q}{e} \frac{EdcT}{\frac{hc}{e}}$ , which with  $\frac{hc}{e} = \Phi_0$  the flux quantum, and through comparison with (37) leads once again to our final proof that

$$\Delta\varphi^{semi} = -\Delta\varphi^{AB}. \quad (40)$$

We note therefore that even in the electric case, the semiclassical phase difference (between two trajectories) picked up due to the Lorentz force (exerted on them) is once again opposite to the electric “Aharonov-Bohm phase” phase picked up by the same trajectories (due to the electric flux that they enclose).

We should point out once again, however, that although the above elementary considerations apply to semiclassical motion of narrow wavepackets, in this paper we have given *a more general understanding of the above opposite sign-relationships* that applies to general (even completely delocalized) states, and that originates from what could be called “generalized Werner & Brill cancellations” (the cancellations that come out of our new solutions, in cases more general than those of Werner & Brill, including both magnetic and electric fields, both static and  $t$ -dependent situations, and both semiclassical and spread-out quantum states).

In a slightly different vein, the cancellations that we found above give an explanation of why certain classical arguments (invoking the past  $t$ -dependent history of an experimental setup) seem to be successful in giving at the end an explanation of Aharonov-Bohm effects (namely a phase consistent with that of a static Aharonov-Bohm configuration). However, there is again an opposite sign that seems to have been largely unnoticed in such arguments as well (i.e. in Silverman[14], where in his eq.(1.34) there should be an extra minus sign). Our above observation essentially describes the fact that, *if we had actually used a  $t$ -dependent magnetic flux* (with its final value being the actual value of our static flux), then the induced electric field (viewed now as a nonlocal term of the present work) would have cancelled the static Aharonov-Bohm phase. Of course now, this  $t$ -dependent experimental set up has not been used (the flux is static) and we obtain the usual magnetic Aharonov-Bohm phase, but the above argument (of a “potential experiment” that *could have been carried out*) takes the

“mystery” away of why such history-based arguments generally work – although *they have to be corrected with a sign*. The above also gives a rather natural account of the “dynamical nonlocality” character[2] attributed to the various Aharonov-Bohm phenomena (magnetic, electric or combined), although – in the present work – this dynamical quantum nonlocality seems to simultaneously respect Causality, as will be seen in the next Section. This is a rather pleasing characteristic of this theory that, as far as we are aware, has no parallel in other formulations.

## V. FULL (X,Y,T)-CASE

Finally, let us look at the most general spatially-two-dimensional and time-dependent case. This combines effects of (perpendicular) magnetic fields (which, if present only in physically-inaccessible regions, can have Aharonov-Bohm consequences) with the temporal nonlocalities of electric fields (parallel to the plane) found in previous Sections. By working again in Cartesian spatial coordinates, we now have to deal with the full system of PDEs

$$\frac{\partial \Lambda(x, y, t)}{\partial x} = A_x(x, y, t), \quad \frac{\partial \Lambda(x, y, t)}{\partial y} = A_y(x, y, t), \quad -\frac{1}{c} \frac{\partial \Lambda(x, y, t)}{\partial t} = \phi(x, y, t). \quad (41)$$

This exercise is considerably longer than the previous ones but important to solve, in order to see in what manner the solutions of this system are able to *combine* the spatial and temporal nonlocal effects found above. There are now  $3!=6$  alternative integration routes to follow for solving this system (and, in addition to this, the results in intermediate steps tend to proliferate). The corresponding (rather long) procedure for solving the system (41) is described in detail in [5], and 2 out of the 12 solutions that can be derived turn out to be the most crucial for the discussion that will follow. First, by following steps similar to the ones of Section IV, the following temporal generalization of (21) is obtained

$$\begin{aligned} \Lambda(x, y, t) = & \Lambda(x_0, y_0, t) + \int_{x_0}^x A_x(x', y_0, t) dx' + \int_{y_0}^y A_y(x, y', t) dy' + \\ & + \left\{ - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t) + G(y, t) \right\} + f(x_0, t) \end{aligned} \quad (42)$$

with  $G(y, t)$  such that  $\left\{ - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t) + G(y, t) \right\} :$  is independent of  $y$ ,

and from this point on, the third equation of the system (41) is getting involved to determine the nontrivial effect of scalar potentials on  $G(y, t)$ . Indeed, by combining it with (42) there results a wealth of patterns, one of them leading finally to our first solution, namely

$$\begin{aligned} \Lambda(x, y, t) = & \Lambda(x_0, y_0, t_0) + \int_{x_0}^x A_x(x', y_0, t) dx' + \int_{y_0}^y A_y(x, y', t) dy' - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t) + G(y, t_0) - \\ & - c \int_{t_0}^t \phi(x_0, y_0, t') dt' + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y, t') + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x_0, y', t') + F(x, y) + f(x_0, t_0) \end{aligned} \quad (43)$$

with the functions  $G(y, t_0)$  and  $F(x, y)$  to be chosen in such a way as to satisfy the following 3 independent conditions:

$$\left\{ G(y, t_0) - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t_0) \right\} : \text{ is independent of } y, \quad (44)$$

which is of course a special case of the condition on  $G(y, t)$  above (see after (42)) applied at  $t = t_0$ , and the other 2 turn out to be of the form

$$\left\{ F(x, y) + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y, t') \right\} : \text{ is independent of } x, \quad (45)$$

$$\left\{ F(x, y) + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x, y', t') \right\} : \text{ is independent of } y. \quad (46)$$

(It is probably important to note that for the above results the Faraday's law is crucial[5]). As for the constant quantity  $f(x_0, t_0)$  appearing in (43), it again describes possible effects of multiple-connectivity at the instant  $t_0$  (which are absent for simple-connected spacetimes, but will be crucial in the discussion of the van Kampen thought-experiment to be discussed later below).



Eq. (43) is our first solution. It is now crucial to note that an alternative form of solution (with the functions  $G$  and  $F$  satisfying the *same* conditions as above) can be derived, and it turns out to be

$$\begin{aligned} \Lambda(x, y, t) = & \Lambda(x_0, y_0, t_0) + \int_{x_0}^x A_x(x', y_0, t) dx' + \int_{y_0}^y A_y(x, y', t) dy' - \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t_0) + G(y, t_0) - \\ & - c \int_{t_0}^t \phi(x_0, y_0, t') dt' + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y_0, t') + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x, y', t') + F(x, y) + f(x_0, t_0). \end{aligned} \quad (47)$$

In this alternative solution we note that, in comparison with (43), the line-integrals of  $\mathbf{E}$  have changed to the *other* alternative “path” (note the difference in the placement of the coordinates of the initial point  $(x_0, y_0)$  in the arguments of  $E_x$  and  $E_y$  compared to (43)) and they happen to have the same sense as the  $\mathbf{A}$ -integrals, while simultaneously the magnetic flux difference shows up with its value at the initial time  $t_0$  rather than at  $t$ . This alternative form will be shown to be useful in cases where we want to directly compare physical situations in the present (at time  $t$ ) and in the past (at time  $t_0$ ), and the above noted change of sense of  $\mathbf{E}$ -integrals (compared to (43)) will be crucial in the discussion that follows in the next Section. (It is also important here to note that, in the form (47), the electric fields have already incorporated the effect of radiated  $B_z$ -fields in space (through the Maxwell’s equations, see [5]), and this is why at the end only the  $B_z$  at  $t_0$  appears explicitly).

Once again the reader can directly verify that (43) or (47) indeed satisfy the basic input system (41). (This verification is considerably more tedious than the earlier ones but rather straightforward).

But a last mathematical step remains: in order to discuss the van Kampen case, namely an enclosed (and physically inaccessible) magnetic flux (which however is *time-dependent*), it is important to have the analogous forms through a reverse route of integrations (see [5]), where at the end we will have the reverse “path” of  $\mathbf{A}$ -integrals (so that by taking the *difference* of the resulting solution and the above solution (43) (or (47)) will lead to the *closed* line integral of  $\mathbf{A}$ , which will then be immediately related to van Kampen’s magnetic flux (at the instant  $t$ )). By following then the reverse route, and by applying a similar

strategy at every intermediate step, we finally obtain the following solution (the spatially “dual” of (43)), namely

$$\begin{aligned} \Lambda(x, y, t) = & \Lambda(x_0, y_0, t_0) + \int_{x_0}^x A_x(x', y, t) dx' + \int_{y_0}^y A_y(x_0, y', t) dy' + \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t) + \hat{G}(x, t_0) - \\ & - c \int_{t_0}^t \phi(x_0, y_0, t') dt' + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y_0, t') + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x, y', t') + F(x, y) + \hat{h}(y_0, t_0) \end{aligned} \quad (48)$$

with the functions  $\hat{G}(x, t_0)$  and  $F(x, y)$  to be chosen in such a way as to satisfy the following 3 independent conditions:

$$\left\{ \hat{G}(x, t_0) + \int_{y_0}^y dy' \int_{x_0}^x dx' B_z(x', y', t_0) \right\} : \text{ is independent of } x, \quad (49)$$

$$\left\{ F(x, y) + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y, t') \right\} : \text{ is independent of } x, \quad (50)$$

$$\left\{ F(x, y) + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x, y', t') \right\} : \text{ is independent of } y, \quad (51)$$

where again for the above results the Faraday’s law is crucial. The corresponding analog of the alternative form (47) (where  $B_z$  appears at  $t_0$ ) is more important for the discussion that follows and turns out to be

$$\begin{aligned} \Lambda(x, y, t) = & \Lambda(x_0, y_0, t_0) + \int_{x_0}^x A_x(x', y, t) dx' + \int_{y_0}^y A_y(x_0, y', t) dy' + \int_{x_0}^x dx' \int_{y_0}^y dy' B_z(x', y', t_0) + \hat{G}(x, t_0) - \\ & - c \int_{t_0}^t \phi(x_0, y_0, t') dt' + c \int_{t_0}^t dt' \int_{x_0}^x dx' E_x(x', y, t') + c \int_{t_0}^t dt' \int_{y_0}^y dy' E_y(x_0, y', t') + F(x, y) + \hat{h}(y_0, t_0) \end{aligned} \quad (52)$$

with  $\hat{G}(x, t_0)$  and  $F(x, y)$  following the same 3 conditions above. The constant term  $\hat{h}(y_0, t_0)$  again describes possible multiplicities at the instant  $t_0$ ; it is absent for simple-connected spacetimes, but will be crucial in the discussion of the van Kampen thought-experiment.

In (48) (and in (52)), note the “alternative paths” (compared to solution (43) (and (47))) of line integrals of  $\mathbf{A}$ ’s (or of  $\mathbf{E}$ ’s). But the most crucial element for what follows is the need to *exclusively* use the forms (47) and (52) (where  $B_z$  only appears at  $t_0$ ), and the fact that, within each solution, the sense of  $\mathbf{A}$ -integrals is the *same* as the sense of the  $\mathbf{E}$ -integrals. (This is *not* true in the other solutions where  $B_z(\dots, t)$  appears, as the reader can directly see). These facts will be crucial to the discussion that follows, which briefly addresses the so called “van Kampen paradox”.

In [15] van Kampen considered a magnetic Aharonov-Bohm setup, but with an inaccessible magnetic flux that is  $t$ -dependent: he envisaged turning on the flux very late, or equivalently, observing the interference of the two wavepackets on a distant screen very early, earlier than the time it takes light to travel the distance to the screen (i.e.  $t < \frac{R}{c}$ ), hence using the (instantaneous nature of the) Aharonov-Bohm phase to transmit information (on the presence of a confined flux somewhere in space) *superluminally*. Indeed, the Aharonov-Bohm phase at any  $t$  is determined by differences of  $\frac{q}{\hbar c}\Lambda(\mathbf{r}, t)$  with  $\Lambda(\mathbf{r}, t) \sim \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}'$  (basically a special case of (5)). However, if we use, instead, our results (47) and (52) above (that contain the additional nonlocal terms), it turns out that, for a spatially-confined magnetic flux  $\Phi(t)$ , the functions  $G$ ,  $\hat{G}$  and  $F$  in the above solutions can then all be taken zero: **(i)** their conditions are all satisfied for a flux  $\Phi(t)$  that is not spatially-extended (hence, from the 2 conditions on  $G$  and  $\hat{G}$  (eq. (44) and (49)) we obtain  $G = \hat{G} = 0$  since the integrals in brackets are all independent of  $x$  and  $y$ ), and **(ii)** for  $t < \frac{R}{c}$ , the integrals of  $E_x$  and  $E_y$  in the corresponding conditions (eq.(45) and (46)) are already independent of *both*  $x$  and  $y$  (since  $E_x(x, y, t') = E_y(x, y, t') = 0$  for all  $t' < t < \frac{R}{c}$ , with  $(x, y)$  the observation point [the essential point being that at instant  $t$ , the  $\mathbf{E}$ -field has not yet reached the spatial point  $(x, y)$  of the screen, and therefore all integrations of  $E_x$  and  $E_y$  with respect to  $x'$  and  $y'$  will be contributing only up to a light-cone (see Fig.4) and they will therefore give results that are *independent of the integration upper limits*  $x$  and  $y$  – basically a generalization of the striped cases that we saw earlier but now to the case of 3 spatio-temporal variables (with now the spatial point  $(x, y)$  being outside the light-cone defined by  $t$  (see Fig.4; in this figure

the initial spatial point  $(x_0, y_0)$ , taken for simplicity at  $(0, 0)$ , has been supposed to be in the area of the inaccessible flux  $\Phi(t)$ , so that, for  $\sqrt{(x - x_0)^2 + (y - y_0)^2} = \sqrt{x^2 + y^2} = R$ , we have indeed that  $ct < R$ , as written on the figure)); we therefore rigorously obtain  $F = 0$ ). Moreover, the Aharonov-Bohm multiplicities (at  $t_0$ ) lead to cancellation of the  $B_z$ -terms (always at  $t_0$ ), with the final result (after subtraction of the 2 solutions) being

$$\begin{aligned} \Delta\Lambda(x, y, t) = & \int_{x_0}^x A_x(x', y_0, t)dx' + \int_{y_0}^y A_y(x, y', t)dy' - \int_{x_0}^x A_x(x', y, t)dx' - \int_{y_0}^y A_y(x_0, y', t)dy' + \\ & + c \int_{t_0}^t dt' \left\{ \int_{x_0}^x dx' E_x(x', y_0, t') + \int_{y_0}^y dy' E_y(x, y', t') - \int_{x_0}^x dx' E_x(x', y, t') - \int_{y_0}^y dy' E_y(x_0, y', t') \right\}. \end{aligned} \quad (53)$$

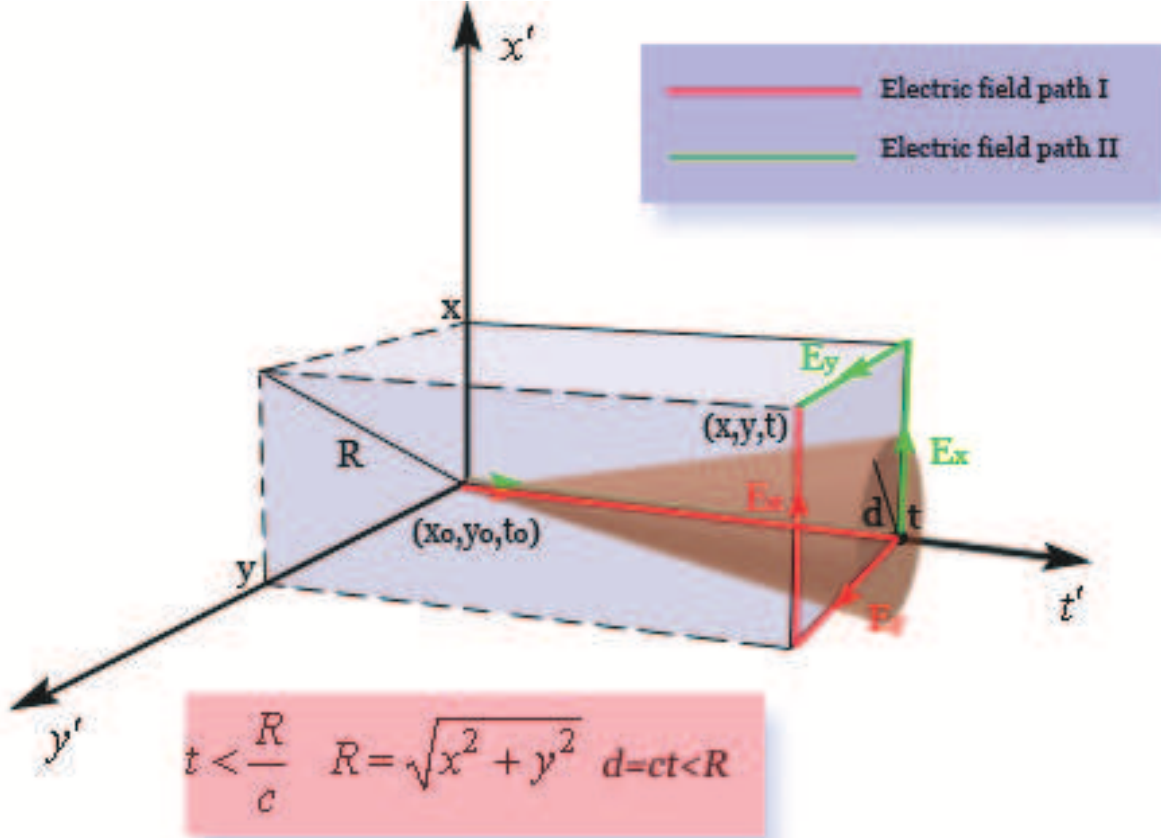
This can equivalently be written as

$$\Delta\Lambda(x, y, t) = \oint \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}' + c \int_{t_0}^t dt' \oint \mathbf{E}(\mathbf{r}', t') \cdot d\mathbf{r}' \quad (54)$$

which, with  $\oint \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}' = \Phi(t)$  the instantaneous enclosed magnetic flux and with the help of Faraday's law  $\oint \mathbf{E}(\mathbf{r}', t') \cdot d\mathbf{r}' = -\frac{1}{c} \frac{d\Phi(t')}{dt'}$ , gives

$$\Delta\Lambda(x, y, t) = \Phi(t) - (\Phi(t) - \Phi(t_0)) = \Phi(t_0). \quad (55)$$

Although  $\Delta\Lambda$  is generally  $t$ -dependent, we obtain the intuitive (causal) result that, for  $t < \frac{L}{c}$  (i.e. if the physical information has not yet reached the screen), the phase-difference turns out to be  $t$ -independent, and leads to the magnetic Aharonov-Bohm phase that we *would* observe at  $t_0$ . *The new nonlocal terms have conspired in such a way as to exactly cancel the Causality-violating Aharonov-Bohm phase* (that would be proportional to the instantaneous  $\Phi(t)$ ). This gives an honest resolution of the van Kampen “paradox” within a canonical formulation, without using any vague electric Aharonov-Bohm argument (as there is no multiple-connectivity in the  $(x, t)$ -plane[9]). An additional physical element is that, for the above cancellation, it is not only the  $E$ -fields but also the  $t$ -propagation of the  $B_z$ -fields (the full “radiation field”) that plays a role[5].



Use of other 10 solutions that rigorously come out from the basic system of PDEs can also address bound-state analogs (i.e. in  $t$ -dependent magnetic flux-driven 1-D nanorings, such as in [16]) or even “electric” analogs of the above van Kampen case: In Peshkin’s review[3] on the electric Aharonov-Bohm effect, the author correctly states “One cannot wait for the electron to pass and only later switch on the field to cause a physical effect”. Although Peshkin uses his eq.(B.5) and (B.6) (based on the incorrect (5)), he carefully states that it is not the full solution; actually, if we view it as an *ansatz*, then it is understandable why he needs to enforce a *condition* (his eq.(B.8), and later (B.9)) on the electric field outside the cages (in order for certain (annoying) terms (resulting from a minimal substitution due to the incorrect ansatz) to vanish and for (B.5) to be a solution). But then he notes that the extra condition cannot always be satisfied (hence (B.5) is not really the solution for all times), drawing from this the above qualitatively correct conclusion on Causality. As it turns out, our treatment gives exactly what Peshkin describes in words (with the total “radiation field” outside the cages being once again crucial in recovering Causality), in a

similar way as in the case presented above in this Section for the usual (magnetic) version of the van Kampen experiment. In this “electric analog” that we are discussing now, the causally-offending part of the electric Aharonov-Bohm phase difference will be cancelled by a magnetic type of phase, that originates from the magnetic field that is associated with the  $t$ -dependence of the electric field  $\mathbf{E}$  outside the cages. We conclude that our (exact) results accomplish precisely what Peshkin has in mind in his discussion (on Causality), but in a direct and fully quantitative manner, and with *no ansatz* based on an incorrect form.

## VI. DISCUSSION

Returning for the moment to only one of the many misconceptions briefly pointed out in this paper, we should emphasize further that improper uses of simple Dirac phases in the literature are not rare or marginal: even in Feynman[11] it is stated that the simple phase factor  $\int^x \mathbf{A} \cdot d\mathbf{r}' - c \int^t \phi dt'$  is valid even for dynamic fields; this is also explicitly stated in Erlichson’s review[17] – Silverman[14] being the only report that we are aware of with a careful wording about (5) being only restrictedly valid (for  $t$ -independent  $\mathbf{A}$  and  $\mathbf{r}$ -independent  $\phi$ ), although even there the nonlocal terms have been missed.

With respect to the presence of *fields* in the *phases* of quantum mechanical wavefunctions that we find, it should be stressed that at the level of the basic Lagrangian  $L(\mathbf{r}, \mathbf{v}, t) = \frac{1}{2}m\mathbf{v}^2 + \frac{q}{c}\mathbf{v} \cdot \mathbf{A}(\mathbf{r}, t) - q\phi(\mathbf{r}, t)$  there are no fields present, and the view holds in the literature[18] that electric fields or magnetic fields cannot contribute directly to the phase. This view originates from the path-integral treatments widely used (where the Lagrangian determines directly the phases of Propagators), but, nevertheless, our canonical treatment shows that fields *do* contribute nonlocally, and they are actually crucial in recovering Relativistic Causality. Moreover, path-integral discussions[19] of the van Kampen case use wave (retarded)-solutions for  $\mathbf{A}$  (hence in the Lorenz gauge) and are incomplete; our results take advantage of the retardation of *fields*  $\mathbf{E}$  and  $\mathbf{B}$  (true in *any* gauge), and *not* of potentials. In addition, Troudet[19] correctly states that his path-integral treatment is good for not highly-delocalized states in space, and that in case of delocalization the proper treatment “would be much more complicated, and would require a much more complete analysis”. It is fair to state that such a complete analysis has actually been provided in

the present work. It should be added that the van Kampen “paradox” seems to be still thought of as “remarkable” [20]. The present work has provided a natural and general *resolution*, and most importantly, through nonlocal (and Relativistically causal) propagation of wavefunction-phases.

Finally, in trying to explore an even broader significance of the new solutions, one may wonder about possible consequences of the nonlocal terms if these are included in more general physical models that have a gauge structure (in Condensed Matter or High Energy Physics). It is also worth mentioning that if one follows the same “unconventional” method (of solution of PDEs) with the Maxwell’s equations for the electric and magnetic fields (rather than with the PDEs of eq.(2) for the potentials that give  $\Lambda$ ), the corresponding nonlocal terms can be derived in a similar manner, and one can then see that these nonlocal terms essentially demonstrate the causal propagation of the radiation electric and magnetic fields outside physically inaccessible confined sources (i.e. solenoids or electric cages). Although this is of course widely known at the level of classical fields, a major conclusion that can be drawn from the present work (at the level of gauge transformations) is that *a corresponding Causality also exists at the level of quantum mechanical phases* as well, and this is enforced by the nonlocal terms in  $t$ -dependent cases. It strongly indicates that the nonlocal terms found in this work at the level of quantum mechanical phases reflect a causal propagation of wavefunction phases **in the Schrödinger Picture** (at least one part of them – the one containing the fields – which competes with the Aharonov-Bohm types of phases containing the potentials). This is an entirely new concept (given the *local nature* but also the *non-relativistic character* of the Schrödinger equation) and deserves to be further explored. It would indeed be worth investigating possible applications of the above results (of nonlocal phases of wavefunctions, solutions of the local Schrödinger equation) to  $t$ -dependent single-*vs* double-slit experiments recently discussed by the group of Aharonov[6] who use a completely different method, with modular variables in the Heisenberg picture (presented as the sole method appropriate for problems of this type). One should also note recent work[7], that rightly emphasizes that Physics cannot currently predict how we dynamically go from the single-slit diffraction pattern to the double-slit diffraction pattern (whether it is in a gradual and causal manner or not) and where a relevant experiment is proposed to decide on (address) exactly this largely unknown issue. Application of our nonlocal terms to such questions in analogous experiments (i.e. by introducing (finite) scalar potentials on slits in a

$t$ -dependent way) provides a completely new formulation for addressing causal issues of this type, and is worth of further investigation. Furthermore, it is worth noting that, if  $E$ 's were substituted by gravitational fields and  $B$ 's by Coriolis force fields arising in non-inertial (rotating) frames of reference, the above nonlocalities (and their apparent causal nature) could possibly have an interesting story to tell about quantum mechanical phase behavior in a Relativistic/Gravitational framework. Finally,  $SU(2)$  generalizations would be an obviously interesting extension of the above  $U(1)$  theory, and such generalizations are rather formally direct and not difficult to make; an immediate physically interesting question would then be whether the new nonlocal terms might have a nontrivial impact on i.e. spin- $\frac{1}{2}$ -states, since these terms would act asymmetrically on opposite spins (the nonlocal  $B_z$ -terms being relevant for Zeeman interactions, and the nonlocal  $E$ -terms possibly having a role if the above results were applied i.e. to Condensed Matter systems with strong spin-orbit coupling[21]).

## VII. CONCLUSIONS

We conclude that a nonlocal and causal behavior exists at the level of quantum mechanical phases, even for solutions of the nonrelativistic (and local) Schrödinger equation and this behavior is enforced by the nonlocal terms derived in the present work (through the well-known causal behavior of fields). Our (exact and analytical) results accomplish precisely what Peshkin has in mind in his discussion (on Causality) of the electric Aharonov-Bohm effect, but in a direct and fully quantitative manner, and with *no ansatz* based on an incorrect form. Another pleasing characteristic of our results that, as far as we are aware, has no parallel in other formulations, is that they give a rather natural account of the “dynamical nonlocality” character[2] attributed to the various Aharonov-Bohm phenomena (magnetic, electric or combined), although – in the present work – this dynamical quantum nonlocality seems to simultaneously respect Causality in a “deterministic way”, i.e. without requiring the necessity of invoking the Uncertainty Principle. The nonlocal terms found in this work at the level of  $\Lambda$  reflect a causal propagation of wavefunction-phases *in the Schrödinger picture*, with possible immediate applications to  $t$ -dependent slit-experiments recently discussed using the Heisenberg picture[6]. Application of our nonlocal terms to such problems (i.e. by introducing  $t$ -dependent scalar potentials on the slits) provides a new



and *direct* formulation for addressing causal issues of such  $t$ -dependent slit-systems. Finally, one cannot refrain from wondering about the analogs of these new nonlocalities in many areas of Physics where geometric or topological phases[22] appear as the central quantities (these always being of the form of integrals of some effective (or emergent) potentials (that are determined by the system, i.e. by band structures, as in the rapidly expanding area of Topological Insulators[21])). A natural thought then arising from the present work is that, if those emerging potentials are not of a type that would correspond to zero fields (in the physically accessible regions), but describe nonvanishing effective fields in (even remote) spacetime regions that are accessible to the particles, then one would expect that the new nonlocalities should be seriously taken into account – these having certain dynamical consequences that may have not received an entirely proper treatment in earlier works. Given the popularity and importance of these areas in the whole of Physics, issues such as the above would certainly deserve further study.

### VIII. ACKNOWLEDGEMENTS

Graduate students Kyriakos Kyriakou and Georgios Konstantinou of the University of Cyprus and Dr. Areg Ghazaryan of Yerevan State University are acknowledged for having carefully reproduced all results. Georgios Konstantinou is also gratefully acknowledged for having drawn all figures of this article. Dr. Cleopatra Christoforou of the Department of Mathematics and Statistics of the University of Cyprus is acknowledged for a discussion concerning the mathematical method followed.

- 
- [1] Y. Aharonov and D. Bohm, “Significance of Electromagnetic Potentials in the Quantum Theory,” *Physical Review*, Vol. 115, No. 3, 1959, pp. 485-491.
  - [2] S. Popescu, “Dynamical quantum nonlocality,” *Nature Physics*, Vol. 6, 2010, pp. 151-153.
  - [3] M. Peshkin & A. Tonomura, “The Aharonov-Bohm Effect,” *Lecture Notes in Physics* Vol. 340, Springer-Verlag, Berlin, 1989: Part I, Appendix B, pp. 27-28.
  - [4] H. R. Brown & P. R. Holland, “The Galilean covariance of quantum mechanics in the case of external fields,” *American Journal of Physics*, Vol. 67, 1999, pp. 204-214.

- [5] K. Mouloupoulos, “Nonlocal phases of local quantum mechanical wavefunctions in static and time-dependent Aharonov-Bohm experiments,” *Journal of Physics A: Mathematical and Theoretical*, Vol. 43, No. 35, 2010, Article ID: 354019, pp. 1-32.
- [6] J. Tollaksen, Y. Aharonov, A. Casher, T. Kaufherr, and S. Nussimov, “Quantum interference experiments, modular variables and weak measurements,” *New Journal of Physics*, Vol. 12, 2010, Article ID: 013023, pp. 1-29.
- [7] G. P. He, “Flexible scheme for measuring experimentally the speed of the response of quantum states to the change of the boundary condition,” [http://arxiv.org/PS\\_cache/arxiv/pdf/0907/0907.1974v2.pdf](http://arxiv.org/PS_cache/arxiv/pdf/0907/0907.1974v2.pdf) (Nov. 2009), pp. 1-6.
- [8] T. T. Wu and C. N. Yang, “Concept of nonintegrable phase factors and global formulation of gauge fields,” *Physical Review D*, Vol. 12, No. 12, 1975, pp. 3845-3857.
- [9] P. D. Noerdlinger, “Elimination of the Electromagnetic Potentials,” *Il Nuovo Cimento*, Vol. 23, 1962, pp. 158-167.
- [10] F. G. Werner & D. R. Brill, “Significance of Electromagnetic Potentials in the Quantum Theory in the Interpretation of Electron Interferometer Fringe Observations,” *Physical Review Letters*, Vol. 4, No. 7, 1960, pp. 344-347.
- [11] R. P. Feynman, R. B. Leighton and M. Sands, “The Feynman Lectures on Physics,” Addison-Wesley, 1964: Vol. II, Chapter 15, p.13.
- [12] B. Felsager, “Geometry, Particles, and Fields”, Springer-Verlag, Berlin 1998: p.55.
- [13] H. Batelaan & A. Tonomura, “The Aharonov-Bohm Effects: Variations on a subtle theme,” *Physics Today*, Vol. 62, No. 9, 2009, pp. 38-43.
- [14] M. P. Silverman, “Quantum Superposition,” Springer-Verlag, Berlin, 2008: p. 13 and p. 19.
- [15] N. G. van Kampen, “Can the Aharonov-Bohm effect transmit signals faster than light?,” *Physics Letters*, Vol. 106A, 1984, pp. 5-6.
- [16] P. G. Luan & C. S. Tang, “Charged particle motion in a time-dependent flux-driven ring: an exactly solvable model,” *Journal of Physics: Condensed Matter*, Vol. 19, 2007, Article ID: 176224, pp. 1-10.
- [17] H. Erlichson, “Aharonov-Bohm Quantum Effects on Charged Particles in Field-Free Regions,” *American Journal of Physics*, Vol. 38, 1970, pp. 162-173.
- [18] R. A. Brown & D. Home, “Locality and causality in time-dependent Aharonov-Bohm interference,” *Il Nuovo Cimento*, Vol. 107B, 1992, pp. 303-316.

- [19] T. Troudet, “Aharonov-Bohm effect versus causality?,” *Physics Letters*, Vol. 111A, 1985, pp. 274-276.
- [20] H. Lyre, “Aharonov-Bohm Effect,” In: Greenberger M., Hentschel K. and Weinert F., Ed., *Compendium of Quantum Physics*, Springer-Verlag, Berlin, 2009, pp. 1-3.
- [21] M. Z. Hasan & C. L. Kane, “Colloquium: Topological insulators,” *Reviews of Modern Physics*, Vol. 82, No. 4, 2010, pp. 3045-3067.
- [22] A. Shapere & F. Wilczek, “Geometric Phases in Physics,” World Scientific, 1989.

### FIGURE CAPTIONS

**Figure 1.** (Color online): Examples of simple field-configurations (in simple-connected regions), where the nonlocal terms exist and are nontrivial, but can easily be determined: (a) a striped case in 1+1 spacetime, where the electric flux enclosed in the “observation rectangle” is dependent on  $t$  but independent of  $x$ ; (b) a triangular distribution in 2-D space, where the part of the magnetic flux inside the corresponding “observation rectangle” depends on *both*  $x$  and  $y$ . The appropriate choices for the corresponding nonlocal functions  $g(x)$  and  $\hat{g}(t)$  for case (a), or  $g(x)$  and  $h(y)$  for case (b), are given in the text (Sections III and IV respectively).

**Figure 2.** (Color online): The standard double-slit apparatus with an additional strip of a perpendicular magnetic field  $B$  of width  $W$  placed between the slit-region and the observation screen. In the text we deal for simplicity with the case  $W \ll L$ , so that deflections (of the semiclassical trajectories) due to the Lorentz force, shown here for a negative charge  $q$ , are very small.

**Figure 3.** (Color online): The analog of Fig.2 (again for a negative  $q$ ) but with an additional electric field parallel to the observation screen that is turned on for a time interval  $T$ . In the text we deal for simplicity with the case  $T \ll \frac{L}{v}$  (with  $v = \frac{1}{m} \frac{h}{\lambda}$ ,  $\lambda$  the de Broglie wavelength), so that deflections (of the semiclassical trajectories) due to the electric force are again very small. For both Fig.2 and 3, it is shown in the text that  $\Delta\varphi^{\text{semiclassical}} = -\Delta\varphi^{AB}$ , hence we observe an extra minus sign compared to what is usually reported in the literature.

**Figure 4.** (Color online): The analog of paths of Fig.1 but now in 2+1 spacetime for the van Kampen thought-experiment, when the instant of observation  $t$  is so short that the physical information has not yet reached the spatial point of observation  $(x, y)$ . The two solutions (that, for wavepackets, have to be subtracted in order to give the phase difference

at  $(x, y, t)$ ) are given in the text, and are here characterized through their electric field  $E$ -line integral behavior: “electric field path (I)” (the red-arrow route) denotes solution (52), and “electric field path (II)” (the green-arrow route) denotes solution (47). Note that the strips of Fig.1(a) have now given their place to a light-cone. At the point of observation (that lies outside this light-cone) the Aharonov-Bohm phase difference has now become “causal” due to cancellations between the two solutions (the two “electric field paths” above).